Summer 8-1-2021

Development and Initial Evaluation of the Assessment of Teacher Expectations and Practices (ATEP)

Kimberly Barajas

DEVELOPMENT AND INITIAL EVALUATION OF THE ASSESSMENT OF TEACHER EXPECTATIONS AND PRACTICES (ATEP)

by

Kimberly G. Barajas

A Dissertation
Submitted to the Graduate School,
the College of Education and Human Sciences
and the School of Psychology
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Approved by:

Stephanie D. Smith, Ph.D., Committee Chair
Sara S. Jordan, Ph.D.
Randolph Arnau, Ph.D.
Brad Dufrene, Ph.D.

August 2021

Published by the Graduate School

ABSTRACT

Teacher practices and expectations are important factors for students' academic and behavioral functioning (Andersson & Palm, 2016; Oliver, Wehby, & Reschly, 2011; Rimm-Kaufman, Storm, Sawyer, Pianta, & LaParom 2006; Rubie-Davies, 2007; Sandholtz, 2011). The current measures available require a great deal of resources (i.e., time, money, personnel), have poor psychometric properties, or are not comprehensive (e.g., facets assessed; assessment of teacher practices for different grade levels). Given these concerns, the proposed study aimed to develop a psychometrically sound measure that is time and cost efficient and comprehensively assesses the multi-faceted construct of teacher practices. This measure is expected to allow teachers to self-evaluate their teaching practices, identify areas for further development, and track their progress over time. Items for the Assessment of Teacher Practices and Expectations (ATEP) were developed following a thorough review of the extant literature and feedback from experts in education. A total of 269 first through twelfth grade teachers recruited via Qualtrics and social media platforms completed measures used for the present analyses. Exploratory factor analyses supported a five-factor structure and a total of 58 items with high factor loadings from the original 139-item pool were retained. Results also provided good evidence of internal consistency, and some evidence of concurrent and convergent/discriminant validity. In sum, the present study provides promising findings for the ATEP. Future studies should further examine the factor structure of the ATEP using a representative sample of teachers and comparing the performance of the ATEP in assessing teacher practices to well-established classroom observational measures.

ACKNOWLEDGMENTS

I would like to express my deepest and sincerest appreciation to my graduate career mentor, Dr. Stephanie Smith. I will forever cherish the knowledge I gained through the opportunities at USM. Thank you for that! I would also like to thank my committee members, Drs. Sara Jordan, Randy Arnau, and Brad Dufrene. Your insight and feedback were essential throughout this project.

DEDICATION

This project is first and foremost dedicated to my incredible family. Dad, mom, Gabi, Nacho, and Nick: I love you all dearly and can never express the amount of pride and gratitude I feel towards you all. To my dad and mom, you are both inspirations and are the embodiment of hard work, dedication, and perseverance. Thank you for sacrificing comfort (long hours on your feet working that restaurant grind, being far from family, leaving the homeland) and time to ensure we had everything we needed. To my Tio Chuy, Tia Betty, Jesus, Andrea, and Daniela, thank you for bringing me joy in the moments when I was able to sneak away from graduate school. To my dear Koda and Kiki, I will be forever grateful for all the unconditional love, loyalty, and endless cuddles at the end of every long day. Lastly, a dissertation cannot undo all the harm done to marginalized communities (undocumented peoples, members of the LGBTQIA+ community, ethnic and racial minorities), but I use this dedication as a promise to work towards closing the gaps in access to education and medical care, and to continue advocating for minority representation in academia.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I - INTRODUCTION

Numerous studies have examined the impact that teachers' classroom practices and expectations have on student outcomes. Indeed, studies support the link between teacher practices and students' academic success and appropriate classroom behaviors (Andersson & Palm, 2016; Oliver, Wehby, & Reschly, 2011; Rimm-Kaufman, Storm, Sawyer, Pianta, & LaParom 2006; Rubie-Davies, 2007; Sandholtz, 2011). Thus far, studies evaluating teacher practices and expectations tend to rely on methods that are often time consuming and expensive (classroom observations: e.g., Framework of Teaching; Classroom Assessment Scoring System) or on teacher or student-rated measures that have poor psychometric properties (e.g., Classroom Environment Scale, Individualized Classroom Environment Questionnaire). As such, it is important to examine teacher practices and expectations via a measure that is psychometrically sound, time and cost-efficient, and as comprehensive as observational systems. The first aim of the present study is to develop a teacher self-report measure that efficiently and comprehensively assesses teacher practices and classroom expectations to: (a) identify areas of focus that can be improved through teacher interventions to improve student outcomes and (b) measure teacher progress as interventions are implemented. The second aim is to assess initial evidence of the reliability and validity of the developed measure.

A Framework of Teacher Quality

Teacher practices and expectations are considered valuable contributors to the overall construct of teacher quality. According to Goe's (2007) Framework of Teacher Quality (Figure 1.1), teacher qualifications (e.g., education, certification, credentials, teacher test scores, experience), characteristics (e.g., attitudes/beliefs, attributes, self-

efficacy, race, and gender), and practices (e.g., classroom management, organization, instructional delivery, expectations) all contribute to the overarching construct of teacher quality. Notably, teacher effectiveness is a different construct than teacher quality (Wenglinsky, 2002). Teacher effectiveness is conceptualized as gains in students' achievement scores, which is calculated by comparing students' predicted achievement against their actual achievement scores (Kane et al., 2011). Standardized tests yield these achievement scores, but there are limitations to using standardized test scores to assess student achievement. For example, it is not feasible to gather this data for school districts that do not have mandated annual testing (e.g., private schools). Additionally, test-based measures allow for the identification of effective teachers, but test scores do little in the way of providing information on what practices or factors make teachers effective. All in all, there is consensus in the research literature that it is not enough to examine movements in students' standardized test scores (i.e., teacher effectiveness), but rather, all variables of Goe's (2007) framework should be considered. Thus, factors (i.e., teacher practices and expectations) contributing to teacher effectiveness must be evaluated to develop targeted interventions for teachers with the end goal of maximizing student outcomes.

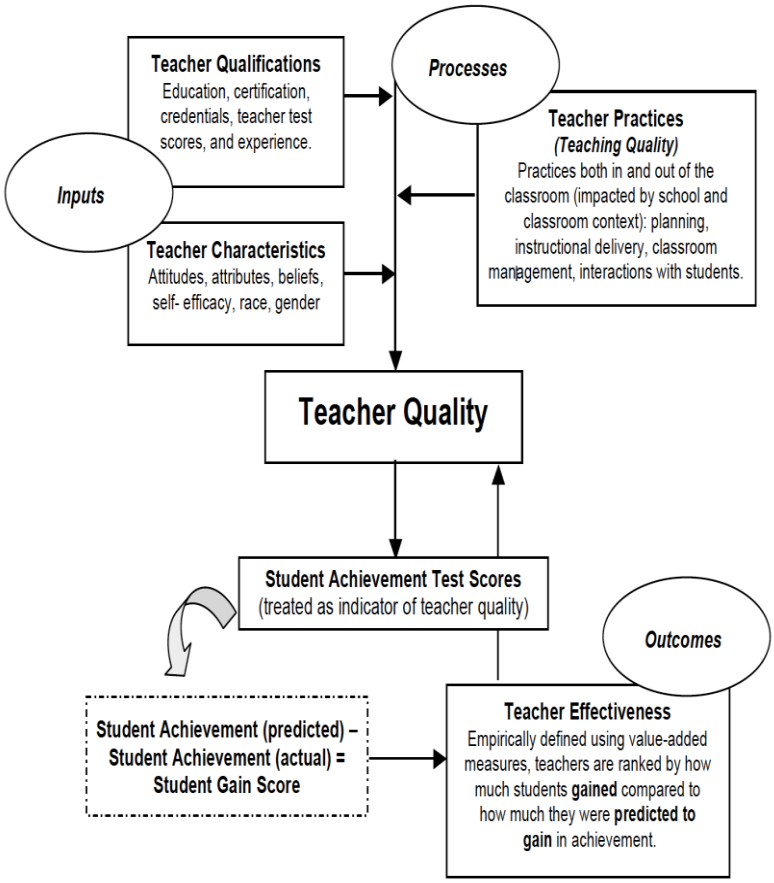

Figure 1.1 *Framework for Teacher Quality according to Goe, 2007.*

Teacher Factors and Student Outcomes

Teacher practices are often viewed as the establishment and maintenance of the learning environment through structure and organization (Brophy, 2006; Doyle, 2006; Duke, 1979; Gettinger & Kohler, 2006; Good & Brophy, 1994; Jones, 1996; Weinstein & Mignano, 1993). Indeed, Good (1979) suggested that teacher practices (e.g., organizational and behavioral management skills) are often effective in differentiating between low- and high-quality teachers. In fact, teacher practices have the greatest impact on students' academic performance when compared to teacher qualifications (e.g., educational attainment) and are comparable to student background in terms of its

influence on student achievement (Wenglinksy, 2002). This notion is further supported by data revealing that teachers who have well-developed practices have a moderate amount of influence on students' basic skills (e.g., foundational reading and math skills), whereas teachers who struggle with teacher practices often have difficulty maintaining student involvement in the class throughout the year (Anderson & Evertson, 1978; Evertson & Anderson, 1978; Weinstein, Romano, & Mignano, 2011; Whitehurst, Chingos, & Lingquist, 2014).

Wang, Haertel, and Walberg (1993) reviewed 28 variables (e.g., positive teacher practices, student-teacher social interactions, peer group, school culture and policies) that were purported to influence student achievement. According to their findings, positive teacher practices had the most influence on student achievement when compared to other variables (e.g., cognitive ability, school demographics). These results suggest that teachers contribute just as much to the students' learning as the students themselves. Indeed, when classrooms are mismanaged and there are no clear teacher practices (e.g., behavior management, organization), there is less time that is available for instruction, which subsequently impacts student learning (Evertson & Emmer, 2017). The following sections define specific subdomains of teacher practices and their importance in establishing a classroom environment that increases students' likelihood of success. These teacher practices have been established as important across grade levels (i.e., elementary through high school; e.g., Brophy, 1988; Doyle, 2006; Evertson & Weinsten, 2006). Additionally, a list of the most important teacher practices presented in the literature can be found in Table 1.2.

Table 1.2 *List of the most important teacher practices currently presented in the literature.*

Important Teacher Practices	References
Establishment and enforcement of rules, routines, procedures, and expectations and the ability restore classroom order	Brophy, 1988; Doyle, 2006; Evertson & Weinstein, 2006; Good & Brophy, 1994; Weinstein & Mignano, 1993
Facilitation of student socialization	Brophy, 1988; Doyle, 1986; Evertson & Weinstein, 2006
Assessment of student performance to provide feedback	Evertson & Emmer, 2017
Organization of resources (e.g., allocation and arrangement)	Brophy, 1988; Doyle, 2006; Henley, 2006; Weinstein & Mignano, 1993
Preservation of classroom pace	Evertson & Emmer, 2017
Maintenance of student attention and the ability to monitor engagement	Brophy, 1988; Doyle, 2006; Evertson & Weinstein, 2006; Good & Brophy, 1994; Jones, 1996; Weinstein & Mignano, 1993

Academic and Behavioral Expectations. Interestingly, when comparing teachers' use of effective versus ineffective teaching strategies, teachers most often use ineffective

teaching strategies for students for whom they have low academic expectations (i.e., fewer learning opportunities, less time devoted to instruction-related activities, exposure to less curricular content, less redirection when distracted, fewer assignments; Proctor, 1984). Although the effect of teacher expectations on academic achievement is smaller than expected (e.g., small, $r = .1$ to $r = .2$ to medium-sized, $d = .43$; Hattie, 2009; Jussim & Harber, 2005), teacher expectations may have a cumulative effect, leading to a marked influence on student achievement over time (e.g., grades over the course of a school year). Indeed, teachers' high academic expectations is linked to greater achievement gains when comparing high versus low expectation students (Rubie, Davies, Hattie & Hamilton, 2006).

Regarding teachers' behavioral expectations for students, clear expectations and feedback is positively related to student achievement, positive classroom environments, and lower rates of punitive efforts (e.g., office discipline referrals, out-of-school suspensions) (Bradshaw, Mitchell, & Leaf, 2010; Little & Akin-Little, 2008). In contrast, poor classroom management has a detrimental impact on student achievement, number of special education referrals, suspensions, and school dropouts (Donovan & Cross, 2002; Harrell, Leavell, van Tassel, & McKee, 2004). Unsurprisingly, teachers who struggle with employing behavior management techniques have difficulty conveying course material to students and are more likely to report symptoms of burnout as well as high levels of stress (Berliner, 1986; Browers & Tomic, 2000; Espin & Yell, 1994). When teachers do not have to constantly re-direct students and continually address problematic behaviors, they subsequently have more instructional time (Marshal, 2001). As such, it is

imperative for teachers to develop clear academic and behavioral expectations for all students.

Knowing the impact of teacher's academic and behavioral expectations is almost as important as teachers knowing how to set and communicate appropriate academic and behaviors expectations. Teachers are able to set high academic expectations and may prevent opportunities for disruptive behaviors by: (a) clearly specifying the course content that will be covered, (b) setting high work standards for all students (e.g., format, neatness, due dates; Evertson et al., 1994; Emmer et al., 2012), (c) devoting a substantial amount of class time to practicing vital learning tasks (Murphy et al., 1982; Emmer et al., 2012; Evertson et al., 1994), (d) describing and demonstrating desired behaviors, (e) providing feedback, and (f) outlining clear consequences (Emmer et al., 2012; Evertson et al., 1994).

Emotional and Social Support. Research suggests that a proactive approach to establishing rules and expectations is not sufficient for maintaining student cooperation and compliance (Emmer et al., 2012; Evertson et al., 1994). Rather, emotional (e.g., recognizing and labeling emotions; understanding emotions; expressing and regulating emotions) and social (encouraging prosocial behaviors and interpersonal effectiveness skills via positive feedback) support within the classroom is essential for supportive teacher-student relationships in children as young as 4 years of age (Jennings & Greenberg, 2009; Hagelskamp, Brackett, Rivers, & Salovey, 2013). Emotional and social support are both important given that supportive teacher-student relationships have been linked to positive academic (increases in scores on standardized tests), socio-emotional, and behavioral outcomes for students (Abbot et al., 1998; Darling-Hammond, Ancess, &

Ort, 2002, Gambone, Klem, & Connell, 2002; Marzano et al., 2013; McNeely, Nonnemaker, & Blum, 2002; Osher et al., 2007). Regarding socio-emotional outcomes, students who perceive having a poor relationship with teachers and low feelings of school connectedness, also perceived themselves as having less academic competence, more delinquent and conduct behaviors, anxiety, and depression when compared to students who self-reported positive relationships with their teachers and high feelings of school connectedness (Murray and Greenberg, 2000). In terms of behavioral outcomes, research has linked students' self-reported perceptions of teacher and school connectedness to lower levels of aggression, alcohol use, and drug use (i.e., cigarettes and marijuana; Resnick et al., 1997). The link between student and teacher relationship quality and students' levels of aggression holds true even when examining peer ratings of both constructs (Hughes, Cavell, and Willson, 2001).

Positive reinforcement (e.g., attention) for appropriate student behavior is needed and is one way to model prosocial interactions and behaviors. Positive feedback allows for a positive and engaging classroom environment that facilitates learning (i.e., instruction time is not lost because of disruptive behaviors) (Conroy, Snyder, Al-Hendawi, and Vo, 2009). Additionally, teachers are incredibly important role-models in their student's lives, as they continually model ways to appropriately regulate emotions in emotionally charged and stressful situations (Pianta et al., 2003). Pianta et al. (2003) proposed that facets of attachment theory (Ainsworth et al., 1978; Bowlby, 1982) also be used in cultivating warm and supportive classrooms via supportive student-teacher relationships that emphasize involvement, trust, and responsiveness.

Assessment. Monitoring student learning is reported to be an important teacher practice given that it allows teachers to assess what is being learned and retained (Hattie, 2003). Classroom assignments (e.g., worksheets, tests) provide opportunities for students to learn and retain information through the application of knowledge and repeated exposure to content (Emmer et al., 2012; Evertson et al., 1994). In addition to assessing student learning, providing consistent, immediate, and specific feedback concerning students' academic performance is essential as it reduces the number of errors that students make in their work and allows teachers to identify students that may require extra assistance or support (Emmer et al., 2012; Evertson et al., 1994).

Organization. Classroom organization (e.g., management of student work; structure of learning activities; classroom pace; classroom routines) is an important skill to assess, given its role in maintaining a classroom environment that is conducive to learning and its positive impact on student achievement (Danielson, 2013; Murphy et al., 1982). Classroom organization allows for less time to be spent in preparation mode (e.g., less time spent getting ready and cleaning up), minimizes distractions and disruptions, and keeps students engaged and attentive (Evertson & Emmer, 2017). When students have difficulty finding class materials, figuring out where to turn in assignments, or finding a seat, it can break up the eb and flow of the classroom; further, this disorganization may cause students to lose interest or focus on the task at hand (Emmer et al., 2012; Evertson et al., 1994; Evertson & Emmer, 2017). Furthermore, optimal classroom pacing (e.g., enough work for students to complete, tasks that require students to begin working immediately, and covering an appropriate amount of material within a class period) is also considered an important aspect of classroom organization and

involves striking a delicate balance between students feeling overwhelmed if too much course content is covered or bored if too little course content is covered. Lastly, classroom routines allow instruction to take place in a predictable, focused, and fluid way, thus reducing the amount of cognitive processing that takes place when there is not a predictable daily sequence (Leinhardt et al., 1987). In sum, classrooms that are rated as having more structure promote appropriate academic and social behaviors (Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008).

Instructional Techniques. Teacher instructional practices include techniques aimed at increasing students': (a) understanding and engagement with educational topics and (b) abilities to engage in higher-order thinking skills (e.g., reflection, critical thinking, problem solving, reasoning skills) (Von Secker and Lissitz, 1999). Studies have found that students are more likely to develop higher-order thinking skills if prompted by their teachers to engage in student-directed learning (e.g., opportunities for individuals to problem-solve on their own; also referred to as active-learning or discovery-oriented instruction) versus teacher-directed learning (e.g., large-group instruction, drills, recitation; also referred to as passive-learning or direct instruction) (Chase & Khlar, 2017; Johnson & Barrett, 2017; Von Secker and Lissitz, 1999). Cooperative group learning, inquiry-based activities, the use of materials and manipulatives that encourage and foster representations of concepts, and the application of course content to real-world problems are all examples of instructional practices that have been suggested to facilitate student learning (Von Secker and Lissitz, 1999). Indeed, an association between teachers' instructional practices and test score gains on standardized tests has been repeatedly

found in the extant literature (Chase & Khlar, 2017; Mayer, 1998; Carpenter et al., 1998; Cobb et al., 1992; Fuson et al., 1997).

Teacher Practices and Classroom Environment

The previously reviewed literature reviewed the influence that teacher practices have on creating a positive classroom environment that subsequently allows students to feel supported and helps facilitate learning (Allen, 2010; Conroy, Sutherland, Snyder, & March, 2008). There are several negative teacher practices (e.g., over-reliance on punitive methods; unclear rules) that are associated with poor classroom environments and can lead to increases in student aggression, poor peer relations, and off-task behaviors (Barth, Dunlap, Dane, Lochman, & Wells, 2004). In contrast, positive classroom environments are related to students' academic self-efficacy, high grades, and increases in motivation (Dorman, 2001; Moos & Moos, 1978). Zins, Weissberg, Wang, and Walberg (2004) reported that interpersonal, instructional, climate, and environmental supports that were associated with improved academic performance include: orderly and safe school and classroom environments, caring relationships between students and teachers, teaching approaches that foster cooperative learning and proactive classroom management (e.g., positive feedback), and adult and peer norms that reflect high expectations and support.

Considering past studies have noted a clear connection between classroom environments and teacher practices (Allen, 2010; Conroy et al., 2008), several frameworks and interventions have been developed with the goal of improving teacher practices that are important for students' academic, socio-emotional, and behavioral outcomes. Examples of such interventions and frameworks include Positive Behavioral Interventions and Supports (PBIS; Sugai and Horner, 2006) and Preventing, Acting

Upon, and Resolving (PAR) Comprehensive Behavior Management System (PAR; Rosenberg and Jackman, 2003). Recent studies are underway examining the impact of interventions and frameworks on teacher practices (Baule & Superior, 2020; Petrasek, Noltemeyer, Green, & Palmer, 2021; Kraft & Blazer, 2013; Wang, Spalding, Odell, Klecka, & Lin, 2010), but a gap in the literature remains: teachers need a cost-effective tool that helps aide them in self-evaluating their functioning in the important areas that are known to contribute to student success. Some school districts have teacher practices reliably and validly assessed through research studies or by being in a well-funded area; however, once these resources are no longer available, teachers need a quick and efficient method of monitoring and changing their own practices.

Existing Measures of Teacher Practices

Currently, there are several existing measures designed to assess teacher practices and expectations. Specifically, existing measures include behavioral observation systems and teacher- and student-rated measures. Observational methods are comprehensive in their coverage of teacher practices but tend to be time intensive and costly. Furthermore, although behavioral observation systems provide both summative and formative data (Little, Bell, & Goe, 2009), a study conducted by Brandt et al. (2007) found that only 8% of districts in their sample used classroom observation data as a tool to improve teachers' skills. In contrast, teacher- and student-rated measures tend to take a narrowband approach in the assessment of teacher practices and oftentimes there is limited data available about their psychometric properties or their evidence of reliability and validity is lacking. The following paragraphs describe these existing measures and review

evidence of reliability and validity, the training required for administration, and associated costs (if applicable).

Observational Systems

The two most widely used and cited behavioral observation systems are the Framework for Teaching (FFT; Danielson, 1996) and the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008). The Framework for Teaching assesses teacher practices that promote student learning; the FFT is appropriate for use across grade levels and subject areas (Brandt, Mathers, Oliva, Brown-Sims, & Hess, 2007; Little et al., 2009). Teachers are rated across four domains: Planning and Preparation (domain 1; e.g., assignments and tasks are clear, aligned with the curriculum, monitor student learning via assessments); Classroom Environment (domain 2; e.g., organized classrooms, maximization of instructional time, classroom procedures, effective use of physical space); Instruction (domain 3; e.g., engage students in learning, provide specific feedback); Professional Responsibilities (domain 4; e.g., improvements in teaching through trainings, supportive of colleagues' ongoing learning). These domains are further broken down into fifteen standards outlining specific teacher practices (e.g., Standard 1.2 states "the teacher uses a variety of assessments that align with appropriate testing standards"). Teachers are rated across the FFT domains and subdomains on a scale that has the following anchors: unsatisfactory, basic, proficient, and distinguished.

There are no known studies that have examined the factor structure of the FFT. However, studies have examined validity and reliability by observing and rating elementary and secondary classrooms in 6 school districts across the United States. Two

external observers from the FFT research team yielded inter-rater agreement, which ranged from 52% to 79% for Domains 1 through 3. Domain 4 was not included in these studies, and thus, there are no studies showing inter-rater reliability for that domain (Milanowski et al., 2011). Regarding evidence of convergent validity, correlations between teachers' FFT ratings and teachers' effectiveness on student learning (represented by the amount of improvement between students' test scores at the beginning of the year versus the end of the year) showed weak to moderate correlations (r's ranging from .1 to .3; Milanowski et al., 2011). Evidence of criterion validity was obtained by Kane et al. (2011) after finding that FFT ratings significantly predicted student achievement.

The FFT is quite costly to administer with measurement materials costing over $100 to get started and each observational form costing $20. Further, two-day training workshops range from $700 - $1,250 to obtain certification. For administration purposes, it is suggested that two raters code during one full lesson and that three additional, shorter observations are made following that initial session. Raters are expected to become re-certified every three years. In sum, although the FFT is comprehensive in assessing teacher practices, it might not be a realistic tool for all school districts when considering the amount of resources required to use it.

The second observation system is the Classroom Assessment Scoring System (CLASS) and was developed by Pianta, La Paro, and Hamre (2008). The CLASS is an observation system used only by trained and certified observers that assesses the quality of instruction and classroom interactions between students and teachers in kindergarten to fifth grade (Pianta et al., 2008). The CLASS is often used to monitor classroom quality

(i.e., classroom organization, emotional support, instructional support) for accountability purposes (i.e., ensuring teachers are engaging in effective practices; Mantzicopoulos, French, Patrick, Watson, & Ahn, 2018). The quality of the teacher and student interactions are assessed according to 3 domains: *Emotional Support* (i.e., Positive Climate: degree of mutual emotional connection, respect, and expression of positive emotions between teachers and students; Negative Climate: extent of negative emotions between teachers and students; Teacher Sensitivity: teachers' awareness and responsiveness to children's academic and emotional needs; Regard for Student Perspectives: consideration given to students' interests, motivation, and points of view), *Classroom Organization* (i.e., Behavior Management: monitoring, preventing, redirecting children's behaviors; Productivity: efficiency in organizing routines, activities, and instruction; and Instructional Learning Formats: support for student learning, interesting material, active participation, maximum learning opportunities), and *Instructional Support* (i.e., Concept Development: promotion of higher-order thinking skills; Quality of Feedback; and Language Modeling: support and encouragement of children's language). The three overarching domains (i.e., Emotional Support, Classroom Organization, and Instructional Support) have been corroborated by studies using confirmatory factor analyses (Pianta et al., 2008). Notably, an abbreviated version of the CLASS exists and is entitled the Individualized Classroom Assessment Scoring System (inCLASS). The inCLASS was designed to be an observational tool that examines preschool students' interactions with their tasks, peers, and teachers (Downer, Booren, Lima, Luckner, & Pianta, 2010). Because the present study is focused on measuring

teacher practices for 1st through 12th grade classrooms, the inCLASS will not be reviewed here.

The CLASS' technical manual outlines six studies that report psychometric information for the CLASS (Pianta, La Paro, & Hamre, 2008). There is evidence of criterion validity such that the CLASS domain scores predict student academic achievement and behavioral outcomes (Howes et al., 2008; Mashburn et al., 2008). Specifically, classroom quality (total score on the CLASS) was associated with children's academic performance at the end of the school year. This relationship was still significant, even after controlling for several covariates (e.g., maternal education, ethnicity, and gender). Instructional Support was the domain that significantly and consistently predicted student achievement across studies. Additionally, the Emotional Support domain was associated with increases in children's expressive and receptive language scores, increases in social competence, and decreases in behavior problems (Howes et al., 2008; Mashburn et al., 2008). Lastly, the Classroom Organization domain was associated with children's classroom engagement, self-control, and literacy gains (Ponitz, Rimm-Kaufman, Brock, & Nathanson, 2009; Rimm-Kaufman, Curby, Grimm, Nathanson, & Brock, 2009).

Correlations between the CLASS' three domains ranged from .11 to .79 in a sample consisting of 694 pre-school and 730 kindergarten classrooms (Hamre, Pianta, Mashburn, & Downer, 2007). In another study using the most recent version of the CLASS and using a sample of 164 kindergarten classrooms, the correlations between the three domains were found to range from .77 to .89 (Hamre, Justice, Pianta, Kilday, Sweeney, Downer, & Leach, 2010), suggesting that these domains are assessing similar

constructs. The average inter-rater agreement was reported to be 87% and ranged from 78% to 96%. The CLASS generally takes about eighty minutes to administer: this time is broken up into four "segments" that are conducted throughout one day and consist of 10 minutes of observation and 10 minutes of coding. Notably, there is also an observation system for grades 6-12, but little information is provided regarding its psychometric properties. Training for the CLASS is time-consuming (two-day training) and costly (i.e., approximately $670 per person; manual costs $49.50; and $25.00 for 10 scoring forms).

In sum, classroom observations are considered the most direct way to evaluate teacher practices and can be used to monitor teacher progress and identify areas of improvement (i.e., formative) and can also be used to compare information against specific benchmarks (i.e., summative) (Little et al., 2009). However, major barriers decrease the likelihood that school districts are able to use these observation systems. Classroom observations require a great deal of resources that not all school districts have access to. Specifically, they are often time intensive for both training of raters and its administration in the classroom and are also expensive because of the extensive training required, as well the cost for administration and scoring. Additionally, in considering the discrepancies between correlations found between the three domains, caution should be used in classifying these systems as the gold standards.

Teacher-, Student-, and Observer-Rated Measures

Besides systematic behavioral observation methods, teacher- and student-report measures provide an alternative method of evaluating teacher practices (Little et al., 2009). In general, teacher-, student-, and observer report measures are versatile in that they can provide formative or summative data (Little et al., 2009). Although most

measures assessing teacher practices are designed for either teacher and student raters, the Classroom Strategies Scale (CSS; Reddy, Fabiano, & Dudek, 2013) has both an observer rating form (CSS-Observer Form) and a teacher rating form (CSS-TF). According to Reddy et al. (2013), the two-informant approach facilitates conversations across raters so areas of strengths or weaknesses missed through self-assessment may be captured by an outside observer.

The CSS-OF and CSS-TF was developed for kindergarten to 5th grade teachers. The CSS-OF consists of three parts, with one part focused on the frequency count of specific teaching strategies used by instructors and the other part focused on the presence of certain classroom items/procedures. The third part of the CSS-OF is a rating scale comprised of two primary scales: Instructional Strategies (IS; 26 items) and Behavioral Management Strategies Rating Scale (BMS; 23 items). The CSS-TF essentially adapted the items from the two subscales entitled Instructional Strategies and Behavioral Management Strategies Rating Scale from the CSS-OF so they could be completed by teachers. These two subscales of the CSS-OF and the CSS-TF are rated on a 7-point Likert scale, ranging from "never used" (1) to "always used" (7). Overall, psychometric analyses for the CSS-OF and TF provide evidence of reliability and validity. Specifically, the CSS-TF was found to have good internal consistency; Cronbach's alphas were estimated to be .93 and .94 for the IS and BMS scales, respectively (Reddy et al., 2015). Evidence of validity has not been examined.

The Classroom Environment Scale (CES; Fisher & Fraser, 1983) has two rating forms for teachers and students and was developed for teachers and students in middle and high school classrooms. It asks respondents to provide information on their current

classroom environment ("Actual") and their preferred classroom environment ("Preferred"; Fisher & Fraser, 1983). The CES contains 10 true/false items that are presented for each of its 9 subscales: involvement (e.g., student involvement, e.g., "students daydream a lot in this class"), affiliation (e.g., "students in this class get to know each other really well"), teacher support (e.g., "the teacher takes a personal interest in the students"), task orientation ("the teacher often takes time out from the lesson plan to talk about other things"), competition (e.g., "some students always try to see who can answer questions first"), order and organization (e.g., "assignments are usually clear so everyone knows what to do"), rule clarity (e.g., "there is a clear set of rules for students to follow"), teacher control (e.g., "students don't always have to stick to the rules in this class"), and innovation (e.g., "new and different ways of teaching are not tried very often in this class"). Teacher and student forms that assess the actual environment show less than desirable estimates of internal consistency (teacher form: Cronbach alpha's ranging from .57 to .77; student forms: Cronbach alpha's ranging from .56 to .78; Fisher & Fraser, 1983). Further, no evidence of validity was provided in studies examining the psychometric properties of this measure (e.g., Fisher & Fraser, 1983; Shochet & Smith, 2014). The median correlations of each scale with the other eight scales ranged from .13 to .33 (Fisher & Fraser, 1983). The CES is a copyrighted measure so expenses do exist for the manual and report forms, with estimated costs approximating $250 to get started.

Another measure that also provides teacher- and student-report is the Individualized Classroom Environment Questionnaire (ICEQ; Fraser, 1990). The ICEQ was developed to differentiate classrooms that use individualized practices (e.g., discussion-based classes) from classrooms that use lecture style practices in middle and

high-school classes (Fraser, 1990). Much like the CES, there are different survey forms that allow students and teachers to rate the current classroom environment and their preferred classroom environment; specifically, ICEQ – Actual, ICEQ – Perceived, and ICEQ – Instructional Practices. In total, there are 50 items, with 10 items on each of the five subscales: Personalization (e.g., "the teacher considers students' feelings"), Participation (e.g., "the teacher lectures without students asking or answering questions"), Independence (e.g., "students choose their partners for group work"), Investigation (e.g., "students find out the answers to questions and problems from the teacher rather than from investigations"), and Differentiation (e.g., "different students use different books, equipment, and materials"). Internal consistency for the form examining the actual environment was less than adequate for some subscales and adequate for others (i.e., teacher form: Cronbach alpha's ranged from .75 to .89; student form: Cronbach alpha's ranged from .61 to .79; Fraser, 1980). The mean correlations between each scale and the other scales ranged from .23 to .39 for the teacher form of the actual environment and the mean correlations between each scale and other scales ranged from .01 to .28 for the student form of the actual environment. The ICEQ is not copyrighted and freely available to the public.

The Learning Environment Inventory (LEI; Anderson & Walberg, 1974) is a student-report measure that was developed to measure high school students' degree of agreement/disagreement with respect to whether statements describe their typical school environment. The LEI was later adapted for students between the ages of 8 to 12 years and was renamed My Class Inventory (MCI; Fraser et al., 1982). LEI questions were rephrased for the MCI to better suite younger students' reading levels, and answer

choices were changed to Yes-No responses. The MCI contains five subscales (i.e., cohesiveness; friction; satisfaction; difficulty; and competitiveness; Byrne, Hattie, Fraser, 1986; Sink & Spencer, 2007). Rather than referring to the extant literature for item development, general assumptions about the types of classroom environments that were deemed more desirable or favorable were made. There are two forms available for the MCI student report that allow students to rate the current classroom environment and their preferred classroom environment.

Regarding psychometric properties of the MCI, internal consistency (Cronbach's alpha) for the form examining the Actual environment was reported to range between .58 to .81 (Fraser and O'Brien, 1985); additionally, the correlations of each scale with the other four scales ranged from .11 to .31 (Fraser and O'Brien, 1985). Notably, a long and short form of the MCI exists, and the subscales of the short and long forms are highly correlated (ranging from .91 - .97).

A teacher-rated version of the MCI (termed My Classroom Inventory – Short Form for Teachers, TCMI-SF; Sink & Spencer, 2017) was later developed and is appropriate for use in kindergarten through sixth grade classrooms. The TCMI-SF uses a 5-point Likert scale ("1" strongly disagree, "5" strongly agree, "3" neutral) and contains 24-items. The TCMI-SF contains five subscales: Competitiveness (e.g., "Some students always try to outperform their peers"); Peer Relations (e.g., "All students in the class are fond of one another"); Satisfaction (e.g., "The students see the class as fun"); Difficulty (e.g., "The schoolwork is too complicated for the students"); and School Counseling Impact (e.g., "The school counselor helps students feel good about learning in this classroom"). The inter-item correlations of the TCMI-SF ranged from .28 to .87 and

alpha coefficients ranged from .66 to .87. Additionally, a confirmatory factor analysis revealed evidence of construct validity. The MCI and TCMI-SF are freely available to the public.

In sum, many teacher-report measures exist for examining teacher practices. However, there are several limitations (e.g., accessibility, cost, psychometric data) that do not make these options viable for school districts wishing to gather formative or summative data on teacher practices. Although other teacher-report measures have been developed, they are not readily available as they are generally used to monitor treatment outcomes for interventions (i.e., The Incredible Years Classroom Management Program) and are copyrighted making accessibility difficult.

Need for a Teacher-rated Measure of Teacher Practices

There are a variety of notable strengths for systematic behavioral observation systems (i.e., Framework for Teaching and the CLASS). For example, the FFT and the CLASS both tap facets that have been deemed to comprise the construct of teacher practices (e.g., classroom organization, behavior management, instructional techniques; Little et al., 2009). Systematic observations are able to accurately capture a wide range of teacher practices and yield specific information about what practices are well-developed or underdeveloped. Additionally, the FFT and the CLASS observation systems have been found to be significantly predictive of student achievement and have some evidence of reliability and validity (Goe et al., 2008). These observation measures also have a variety of uses; specifically, they may serve as progress measures following teacher development interventions or may instead provide summative information. Despite the strengths associated with classroom observations, there are several weaknesses that are worth

noting. First, the evaluation of these measures' psychometric properties are quite limited with manuals and published studies primarily reporting inter-rater agreement and correlations among subscales, and in some instances, internal consistency and associations with student achievement. The data for these psychometric properties are also quite variable, suggesting that it may not be measuring a unitary construct. Second, these observation protocols require extensive resources (e.g., time, money) so their feasibility and usability by schools is questionable. Given these concerns, it is important for school administrators to have an easily accessible and cost-effective way of monitoring teacher practices to inform targeted teacher development interventions that positively impact student outcomes.

The CSS-TF, CES, ICEQ and TCMI-SF are the only known teacher-rated measures that are easily accessible and that have been used in several studies. However, these teacher-rated measures have less than desirable psychometric properties, cover a narrower scope of teacher practices than observational methods, and are designed to assess teacher practices in a rather limited range with respect to grade level. Thus, there is a need to develop a comprehensive, psychometrically sound teacher-rated measure of teacher practices that may be used in any classroom for any grade level to inform teacher development interventions with the goal of improving student outcomes.

Current Study: Assessment of Teacher Expectations and Practices (ATEP)

The ATEP is a teacher-rated measure designed to provide formative data by identifying teacher and classroom practices that would benefit from modification through targeted interventions. The ATEP is the first known measure developed solely for the purpose of tracking teachers' improvements in their classroom practices. The ATEP was

designed to assess teacher practices in a time and cost-effective manner. Specifically, the ATEP takes less than 20 minutes to complete, will be a freely available resource, and does not require intensive training to administer considering it is a teacher-rated measure with no elaborate coding scheme. Further, self-report methods provide teachers with the opportunity to self-reflect and involve themselves in their own evaluations. The aim of the present study was to develop a measure that is as comprehensive as observational coding systems of teacher practices (e.g., emotional support, classroom organization, instructional support), while also demonstrating initial evidence of reliability/validity that meets the agreed upon standards of measurement development in the extant literature (Carretero-Dios & Pérez, 2007). In order to meet these goals, the following steps were taken: 1) the extant literature was examined in order to develop items that tapped the domains suggested to comprise teacher practices; 2) a panel of experts in education were consulted to assess the content validity of the proposed items; 3) analyses were conducted to examine the inter-correlations among the retained items and structure of the measure via exploratory factor analyses; 4) reliability was assessed by calculating estimates of internal consistency; and 5) convergent, concurrent, and discriminant validity was examined through a series of correlational analyses between scores on the ATEP and measures hypothesized to be related and unrelated to the construct of teacher practices.

It was hypothesized that a five-factor structure would emerge from the exploratory factor analysis. Items were expected to load according to the domains they represented. There were also expectations that there would be some evidence of reliability and validity. Specifically, internal consistency as measured by Cronbach's alpha was expected to be above .80 for subscale and total scores. Evidence of

discriminant validity was examined by using the Marlowe-Crowne Social Desirability Scale – Short Form (MCSDS-SF), which was expected to have no relationship with the ATEP. Regarding evidence of concurrent and convergent validity, it was hypothesized that there would be a positive relationship between the ATEP total score and the Satisfaction subscale of the My Class Inventory – Short Form for Teachers (TCMI-SF) and the total score of the Teachers' Efficacy Beliefs System – Self (TEBS-Self). In contrast, there would be a negative relationship between the ATEP total score and the total score of the Maslach Burnout Inventory – Educators Survey (MBI-ES).

CHAPTER II – METHOD

Participants

The ATEP was administered online via a Qualtrics survey to teachers across the United States. Participants were recruited using two modalities: (a) e-mails sent by Qualtrics and (b) the social media platform, Facebook. Inclusion criteria for this study was as follows: (a) general education teachers who taught core curriculum classes (i.e., Math, Science, English, History) to students in 1st through 12th grade, (b) ability to read and write in English, and (c) resident of the United States. In all, 538 total participants signed the consent form and began the survey. Two hundred ninety participants were excluded for a variety of reasons, including: failing two or more attention checks (N = 112), not meeting eligibility criteria (N = 36), not completing the entirety of the survey (N = 133), and not passing quality assurance checks (N = 9; i.e., matching IP addresses, latitude/longitude outside of the United States, answering questions in a different language). The remaining 269 participants were included in the study sample as their data were found to be reliable and valid. Of the teachers included in the sample, 46.5% (N = 125) specialized in teaching one specific grade whereas 54.5% (N = 144) had experience teaching multiple grades. Teachers reported teaching experience with the follow grades: 30% taught elementary (N = 109; first through fifth grade), 13% taught middle (N = 43; fifth through eighth grade), 6% taught high school (N = 16; ninth through twelfth grade), and 18.6% reported experience teaching across different levels (N = 50; elementary, middle, and high school). Experience working in the field of education ranged from 1 to 45 years (M = 14.51, SD = 9.82).

Participants lived in various regions of the United States with 20.8% living in the Northeast, 30.5% living in the Southeast, 24.9% living in the Midwest, 7.1% living in the Southwest, and 16.7% living in the West. Forty-eight teachers in the sample were between the ages of 20-30 years (8%), 82 were between the ages of 31-40 years (30.5%), 60 were between the ages of 41-50 years (22.3%), 56 were between the ages of 51 – 60 years (20.8%), and 23 were 61 years or older (8.6%). In terms of teachers' gender identity, 75.5% of the sample identified as female, 24.2% identified as male, and 0.4% identified as non-binary. Racial and ethnic identity was such that 84% (N = 226) of participants identified as White, 7.4% identified as Black or African American, 6.3% identified as Latinx/Hispanic, 5.65% identified as Asian, 0.7% identified as Indigenous, and 2.2% identified with another racial or ethnic group that was not listed. The present sample of teachers taught either English (17.9%), Math (14.5%), History (7.1%), Science (8.6%), or more than one core class (50.2%). Additional demographic information (e.g., school type) can be found in Appendix F.

Measures

Assessment of Teacher Expectations and Practices. The ATEP was developed to assess teachers' expectations and practices in the classroom. Five distinct domains were identified as relevant to this construct after a thorough review of the extant literature. The Behavioral and Academic Expectations (BAE) domain consisted of 46 items assessing the extent to which behavioral and academic expectations are communicated by teachers to their students (e.g., "students in this class understand what behavior is expected of them," "instructions for class and homework assignments are clearly outlined for students"). The Organization (ORG) domain contained 25 items measuring classroom

structure, pace, and routines (e.g., “students have access to the materials they need,” “the classroom routine frequently requires adjustment”). The Emotional and Social Support (ESS) domain comprised of 22 items assessing teacher’s proclivity to encourage emotional development by helping their students label and regulate their emotions and foster a positive classroom environment (e.g., “a warm and supportive environment is maintained for students in this class,” “students are treated with respect at all times”). The Assessment (AS) domain consisted of 24 items measuring how teachers facilitate test preparation and study skills and evaluate their students learning of course content (e.g., “test content directly reflects learning objectives,” “grading criteria and objectives for all assignments are clearly communicated to students”). Lastly, the Instructional Techniques (IT) domain presented 22 items to evaluate teacher’s encouragement of higher order thinking skills, student engagement, and development of skills helpful for retaining new material/knowledge (e.g., “students are taught how to summarize the key concepts of book chapters or lessons”). More information about the specific domain definitions are available in Appendix G.

The initial item pool of 139 statements were rated on a likert scale using the following anchors: 1 - *Strongly Disagree,* 2 - *Disagree,* 3 – *Agree,* 4 - *Strongly Agree*. A total of 23 items on the ATEP were worded in a different direction than the other items so they required reverse coding to produce a total score (e.g., “ expectations for student behavior change on a regular basis,” “it is difficult to find activities that keep students

engaged"). Decisions regarding item deletion, psychometric data, and the final structure of the ATEP are discussed in the results section.

My Class Inventory – Short Form for Teachers (TCMI-SF). The TCMI-SF (Sink & Spencer, 2007) is a teacher-rated measure designed to assess two areas: (a) classroom environment in terms of the learning environment, classroom climate, emotional and physical safety, and quality of the relationships among students, and (b) the extent to which the school counselor contributes to the classroom environment (e.g., "the school counselor helps students feel good about learning in the classroom"; Adelman & Taylor, 2002). The TCMI-SF contains 24-items that are rated using a 5-point Likert scale ("1" strongly disagree, "3" neutral, "5" strongly agree). Five scales comprise the TCMI-SF: Satisfaction, Peer Relations, Competitiveness, Difficulty, and School Counselor Impact (Byrne, Hattie, Fraser, 1986; Sink & Spencer, 2007). The TCMI-SF is scored by obtaining a raw score for each scale. The inter-item correlations for each scale of the TCMI-SF ranged from .28 to .87 and Cronbach's alpha ranged from .60 to .87 for each scale in the validation sample. Additionally, a confirmatory factor analysis revealed evidence of its construct validity. For the present study, the TCMI-SF was used to determine whether there is evidence of convergent validity for the proposed measure (ATEP). For the current sample, Cronbach's alpha was as follows for each scale of the TCMI-SF: Satisfaction (a = .87), Peer Relations (a = .82), Competitiveness (a = .60), Difficulty (a = .84), and School Counselor Impact (a = .93).

Teachers' Efficacy Beliefs System - Self (TEBS – Self). The Teachers' Efficacy Beliefs System (TEBS-Self; Dellinger, Bobbett, Olivier, & Ellett, 2008) is a measure designed to assess teachers' individual beliefs as they pertain to their own abilities to

successfully perform instructional tasks within their classrooms. The TEBS-Self contains 30-items that are rated using a four-point scale (1-very weak belief in my capabilities, 2-moderate belief in my capabilities, 3-strong belief in my capabilities, 4-very strong belief in my capabilities; Dellinger et al., 2008). Five scales comprise the TEBS-Self: Communication/Clarification, Management/Climate, Accommodating Individual Differences, Motivation of Students, and Higher Order Thinking Skills. Prior estimates of Cronbach's alpha for the TEBS have ranged between .85 – 87 for the scale scores. The TEBS was included to assess for evidence of concurrent validity for the proposed measure (ATEP). For the current sample, Cronbach's alpha for each scale was as follows: Communication/Clarification (a = .88), Management/Climate (a = .87), Accommodating Individual Differences (a = .88), Motivation of Students (a = .80), and Higher Order Thinking Skills (a = .88).

Maslach Burnout Inventory – Educators Survey (MBI-ES). The Maslach Burnout Inventory – Educators Survey (MBI-ES; Maslach et al., 1996) is a 22-item self-report measure used to assess the level of teachers' burnout. The MBI-ES is divided into three subscales (Emotional Exhaustion, Depersonalization, and Personal Accomplishment). Participants rated the frequency with which they experience symptoms of burnout on a seven-point rating scale ranging from 'never' to 'always.' Greater scores on the MBI-ES indicate greater frequency of burnout symptoms. Cronbach's alpha for the MBI-ES subscales have been found to range from .71 to .90, with an estimate of internal consistency for the total score being .74. In prior studies assessing the psychometric properties of the MBI-ES, test-retest reliabilities for each subscale were assessed one year from the original administration date and ranged from .54 to .60. Significant correlations

were also found between the MBI-ES and the Job Diagnostic Survey, providing evidence of convergent validity. The MBI-ES was included to assess for evidence of concurrent validity for the proposed measure (ATEP). For the current sample, Cronbach's alpha for the Maslach Burnout Inventory – Educators Survey was .93, .81, and .84 for the Emotional Exhaustion, Personal Accomplishment, and Depersonalization scales, respectively.

Marlowe-Crowne Social Desirability Scale – Short Form (MCSDS - SF). The MCSDS - SF is a self-report measure designed to assess the respondent's tendency to present themselves in a favorable light according to their culture's social norms and standards (Crowne and Marlow, 1960; Reynolds, 1982; Perinelli and Gremigni, 2016). The MCSDS-SF has a total of 13-items that require individuals to respond to items as either true or false. There is some evidence of its internal reliability (r = .76), as items were found to correlate with total scores. Furthermore, evidence of concurrent validity was evaluated in past studies by examining correlations between the MCSDS-SF and the Edwards Social Desirability Scale and results revealed correlations of modest strength between the two measures (r = .4; Reynolds, 1982). The MCSDS- SF was included to assess for evidence of discriminant validity. Cronbach's alpha was .65 for the total score of the Marlow-Crowne Social Desirability Scale – Short Form in this sample.

Demographic Questionnaire. The demographic questionnaire was completed by participating teachers and provided descriptive information about the sample. The questionnaire inquired about race/ethnicity, age, marital status, highest level of education, area of specialization, teaching experience (e.g., number of years teaching, previous/current grades and courses taught, size of the classroom), and family income of

teachers. The questionnaire also captured information about the school (e.g., private or public, estimated number of students who received free or reduced lunch, estimated number of total students in each grade) where the teachers taught.

Procedures

All study procedures were approved by the Internal Review Board of the University at which the research team were affiliated before the initiation of data collection (see Appendix A). Qualtrics and social media platforms (e.g., Facebook) were used to recruit participants to complete study measures. Participants recruited through Qualtrics received e-mail invitations from Qualtrics and were invited to participate in exchange for an incentive of their choosing that equated to $6.00. Qualtrics' e-mail invitations targeted users from their own databases who previously expressed interest in receiving invitations for future surveys and whose profiles listed teacher or educator as a career. Participants recruited through social media were entered into a raffle for a chance to win one out of four available $40.00 Amazon e-cards. Participants who signed the study consent form electronically via Qualtrics were presented with a series of screening questions to determine study eligibility (e.g., general education teacher, 1st - 12th grade teacher). If eligibility criteria were met, participants were presented with a series of questionnaires in the following order: Assessment of Teacher Expectations and Practices, My Class Inventory – Short Form for Teachers (TCMI-SF), Teachers' Efficacy Beliefs System-Self (TEBS-S), Maslach Burnout Inventory – Educators Survey (MBI-ES), Marlowe-Crowne Social Desirability Scale (MCSDS), and the demographic questionnaire. The Qualtrics survey always began with administration of the ATEP to increase the likelihood that participants would be alert and attentive and to reduce the

possibility of fatigue when completing the ATEP (Brophy, Jackson, & Crowe, 2009). The order in which measures were presented remained consistent across all participants. To improve the quality of study data, quality assurance items directed participants to endorse a specific answer and were randomly placed among questionnaire items. Participants who failed 2 or more quality assurance items did not receive compensation for study completion and the survey was terminated. All measures are available in Appendices B – F.

CHAPTER III - RESULTS

Content Validity

Content validity was examined by recruiting three experts who were asked to provide feedback about the ATEP's initial item pool and the extent to which the items accurately assessed the content area of interest (measured by the content validity ratio; Frank-Stromberg & Olsen, 2004; Ayre & Scally, 2014). Two experts were professors with doctoral degrees in education and with expertise in curriculum and instruction, assessment and evaluation, and factors that contribute to and hinder student success. The final expert was an experienced classroom teacher who held a bachelor's degree in education. Open-ended questions provided experts with an opportunity to propose feedback on items, and multiple-choice questions (response options of non-essential, useful, and essential) were presented to calculate the item's content validity ratio (CVR). Specifically, experts were asked to answer four multiple choice questions (i.e., "how essential is the above item in examining the overall constructs of teacher practices and expectations?," "how essential is the above item to the domain being assessed?," "to what extent is the above item developmentally appropriate across grade levels (i.e., 1st through 12th)?," "how clearly stated is the above item?") and one open-ended question ("do you have suggestions of how to modify the item (e.g., item clarity, suggested edits to the item)?") for each item. Recommended changes to the items of the ATEP pertained to the wording of certain items to improve their clarity and brevity, but experts did not recommend that items be added or deleted.

The CVR was calculated by taking the proportion of experts who rated an item as "essential" for measuring a particular domain or construct. The values range from -1 to 1,

where a value of -1 indicates that all raters agreed that an item is non-essential and a value of 1 suggests that all raters agreed that an item is essential (Ayre & Scally, 2014). In an effort to create an efficient measure, a total of 28 items with negative CVR values were deleted all at once, leaving 104 items for the remaining analyses. Many of these deleted items were from the originally hypothesized Behavioral and Academic Expectations and Organization scales, but a few deleted items came from the Emotional and Social Support and Assessment scales.

Preliminary Analyses

After data collection, but prior to conducting the proposed analyses, frequencies and minimum and maximum statistics were examined to ensure no system-errors resulted in outliers or other values outside of the expected range. Descriptive statistics (e.g., mean, standard deviation) were also examined to examine the spread of the data. A total of 23 items on the ATEP were reverse coded prior to conducting further analyses. Per recommendations in the literature, item-total correlations were evaluated next, as it allows for quick identification of items that are not correlated with the scale (Tay & Drasgow, 2012). Of the original 139 items, a total of 7 items had negative item-total correlations and were subsequently deleted (57, 63, 81, 104, 105, 109, 124; Tay & Drasgow, 2012) all at once, leaving a total of 132 items on the ATEP.

Assumptions for Exploratory Factor Analysis

Exploratory factor analysis (EFA) is the recommended statistical analysis to determine inter-relations among items and the factor structure of a measure at the scale level (Hayton, Allen, & Scarpello, 2004; Fabrigar et al., 1999; Williams et al., 2010). However, three assumptions must be met prior to conducting an EFA: (1) the data must

show evidence of normality; (2) items comprising the measure must be significantly correlated (i.e., small to moderate correlations), and (3) there must be a sufficient sample size. First, to examine normality, experts recommend that skewness and kurtosis be examined (Field, 2015). For the present data set, data for all administered measures was determined to be normally distributed according to skewness and kurtosis results. Normal Q-Q plots and histograms were also examined, further confirming the normal distribution. Second, item-total correlations were examined to determine the strength of the relationship between items. Item-total correlations ranged from r = .022 to r = .662, which is considered to be in the acceptable range at this point in the analyses (Costello & Osborne, 2005). Lastly, sample size was considered; with respect to sample size, research offers mixed recommendations about how many participants to include (e.g., Arrindell & van der Ende, 1985; Cattell, 1978; Comrey & Lee, 1992; Gorsuch, 1983; Guilford, 1954; Velicer & Fava, 1998). In general, suggestions range from 100 to 250 participants for EFA, and the present sample size meets those specifications. In sum, our study data satisfy the assumptions necessary to conduct an EFA.

Exploratory Factor Analysis

Exploratory factor analysis requires that three basic decision points are made: (1) decide on the number of factors to extract, (2) choose an extraction method and, (3) select a rotation method. There are several approaches available to determine the number of factors to extract (e.g., Dobrin & Owen, 2019; Gaskin & Happell, 2014). The scree test (Cattell, 1966), parallel analysis (Horn, 1965), and the minimum average partial (MAP) rule (Velicer, 1976a) have been described by experts in the field as “accurate and easy to use” (Costello & Osborne, 2005). The scree test provides a visual method (i.e.,

eigenvalues plotted on a line chart) to identify factors that account for the most variance; however, this method is reported to be subjective and ambiguous when there are multiple elbows or no clear elbows that represent where eigenvalues level off signifying how many factors to retain (Hayton, Allen, & Scarpello, 2004). Parallel analysis operates by comparing eigenvalues from the present data set to eigenvalues generated from a Monte-Carlo simulated matrix estimated from random data. Decisions about the number of factors that are retained are decided by comparing eigenvalues from the original data to the mean eigenvalue of the generated data (Hayton et al., 2004). The MAP procedure identifies an upper limit of factors by partitioning the common and unique variance from the correlation matrix and only retains those factors that share common variance. Parallel analysis allows researchers to account for sampling error, while the MAP analysis examines the likelihood that meaningful correlations have been parsed out by considering the effect of removing the remaining eigenvalues (Caron, 2019). Studies comparing the various methods have found that in certain instances (e.g., studies with small sample size; highly correlated items), PA and MAP sometimes under - or over-estimate the number of factors to retain (Hayton et al., 2004). Because of these discrepant findings, Hayton and colleagues (2004) recommend that parallel analysis be used in conjunction with other methods, such as the scree test and MAP. As such, all three methods were used with the goal of finding the best factor fit. Freely available syntax was used to conduct the PA and MAP analyses. The PA method suggested a seven-factor structure based on a principal

component analysis; the MAP method indicated an eight-factor structure; and the scree plot suggested a five-factor structure for our data.

Regarding the extraction and rotation method, a principal axis factoring extraction method was used, as it is the recommended approach when the goal is to determine the fewest number of factors that account for the greatest amount of variance (Gasking et al., 2014). Promax rotation with Kaiser normalization was specifically chosen since promax is the recommended approach when variables are theoretically expected to correlate with one another (Corner, 2009). Additionally, Kaiser normalization was also used as this helps ensure solutions obtain stability across samples (UCLA Statistical Consulting).

Three EFAs were conducted to examine factor structures for the 5-, 7- and 8-factor models previously suggested by the Scree plot, PA, and MAP, respectively. Models somewhat varied in terms of how much variance was explained by each factor solution where the 5-factor solution explained 50% of the variance, the 7-factor solution explained 45% of the variance, and the 8-factor solution explained 43% of the variance. A close analysis of the factor loadings and pattern coefficients were examined to better understand the underlying construct measured by each factor (Myers et al., 2016). The number of items loading onto each factor was also considered with some experts noting that a factor with fewer than three items should be viewed as “weak and unstable” (Costello & Osborne, 2005). As such, the 8-factor model presented concerns because one of the eight factors extracted from this solution only had three items. The 7-factor model was also ultimately excluded because one of the factors from this solution had items that also loaded onto other factors and displayed higher factor loadings for those other factors.

Ultimately, the five-factor solution was retained, as the items loading onto each extracted factor represented a simple structure and appeared to be evaluating the same underlying construct (see Table 2). The items comprising each factor were further refined using an iterative process to delete items with factor loadings of 0.35 or lower and those with eigenvalues less than 1 (Costello & Osborne, 2005). The item-total correlation and Cronbach's alpha of the factor if an item was deleted was also considered as specific items were evaluated for deletion or retainment. This process led to the deletion of an additional 35 items. Considering one of the main goals in developing this measure was to provide educators with a time-efficient method of examining their expectations and practices, the upper limit of items for each factor was determined to be 15 items after carefully ensuring that certain factor statistics (i.e., Cronbach's alpha if item deleted) would not be negatively impacted. To stay under or at this item limit, content validity ratios were examined once more to identify items that were not rated as essential by all three expert raters and to assess the extent to which an item theoretically fit with the other items on that factor. This process led to the removal of an additional 13 items. Of the 104 items in the initial item pool, 58 items were retained and 46 items were deleted. Content validity ratios, means, standard deviations, corrected item-total correlations, and Cronbach's alphas for the initial item pool are included in Appendix J.

An EFA was conducted once more to examine whether item loadings remained stable once these 46 items were removed (Meyers et al., 2013). A closer examination of the item loadings for the five-factor structure of this final EFA revealed some differences from the originally hypothesized 139 item ATEP measure. Factor 1 from the five-factor solution of this final EFA comprised of 12 items, which captured the extent to which

teachers support and foster prosocial behaviors in their students (e.g., appropriately processing and expressing emotions, using effective interpersonal skills); this factor was labeled “Emotional and Social Support.” Factor 2 comprised of 12 items that captured teachers’ ability to communicate behavioral and academic expectations to their students and provide a stable, predictable classroom routine; this factor was labeled “Classroom Expectations and Routines.” Factor 3 comprised of 14 items that captured teachers’ assessment of students’ learning and methods used to build students critical thinking and study skills and was labeled “Assessment and Instructional Learning.” Factor 4 comprised of 8 items that captured teachers’ difficulty with engaging in effective teacher practices and in consistently maintaining and enforcing classroom rules, expectations, and routines; factor 4 was labeled “Classroom Chaos.” Lastly, Factor 5 comprised 10 items that captured teachers’ ability to cultivate and maintain student engagement and a supportive classroom environment; this factor was labeled “Classroom Climate and Instructional Support.” The final five-factor model accounted for 49% of the variance. The final five-factor structure of the ATEP is presented in Table 3.1 and item statistics for the final measure are presented in Appendix J.

Table 3.1 *Pattern matrix representing the five-factor structure of the ATEP.*

Original Factor Loading	1	2	3	4	5
Factor 1 – Emotional and Social Support					

Table 3.1 Continued

ESS	73. Students are encouraged to take the emotional perspective of others (e.g., “what do you think this person is feeling?”).	**0.845**	0.093	-0.038	0.088	-0.196
ESS	76. Students are coached on how to express their feelings to others.	**0.801**	0.041	-0.015	-0.078	-0.075
ESS	75. Students are encouraged to determine the antecedents of other people’s emotions in novel situations or conflicts (“what happened to this person to make them feel this way?”).	**0.793**	0.027	0.003	-0.034	-0.127

Table 3.1 Continued

ESS	72. Students are taught how to label their emotions through reflection (e.g., “I can tell you are frustrated.”).	**0.779**	0.061	0.079	0.013	-0.267
ESS	77. Students are shown how to appropriately express their emotions through modeling (e.g., teacher states, “I get angry when others laugh at me.”).	**0.753**	-0.056	0.016	0.072	0.031
ESS	79. Students are encouraged to explore the most effective way to cope with their strong emotions.	**0.593**	-0.029	0.026	0.006	0.216

Table 3.1 Continued

ESS	91. Students are given examples of how to appropriately and inappropriately solve conflicts.	**0.544**	-0.026	0.073	0.147	0.138
ESS	88. If a student is struggling to make friends or cannot get along with other students, help or resources are provided (e.g., social skills training).	**0.531**	-0.166	0.131	0.033	0.205
ESS	74. Students feelings are normalized and validated (e.g., “Everyone gets angry.”).	**0.515**	0.153	-0.061	-0.089	0.080

Table 3.1 Continued

ESS	78. Students are encouraged to explore appropriate and inappropriate ways of how characters in a novel or story might express their emotions.	**0.504**	-0.041	-0.124	-0.109	0.277
ESS	92. Teacher uses student conflicts as opportunities to model effective problem solving and communication strategies.	**0.460**	-0.122	0.087	0.001	0.217
BAE	46. Students feel motivated to make good classroom decisions.	**0.374**	0.203	-0.052	-0.046	0.172

Table 3.1 Continued

	Factor 2 - Classroom Expectations and Routines					
BAE	24. Behavioral expectations for this class are developmentally appropriate (e.g., raise hand to speak, follow instructions).	-0.007	**0.779**	-0.248	0.038	0.103
BAE	10. Instructions for class and homework assignments are clearly outlined for students.	-0.022	**0.764**	0.187	-0.008	-0.237
BAE	20. The procedures for turning in class assignments or homework remain consistent.	-0.072	**0.670**	0.255	-0.013	-0.220
BAE	21. Classroom rules are clearly stated to students.	-0.031	**0.665**	-0.149	0.122	0.114

Table 3.1 Continued

BAE	26. Students are provided with examples of appropriate behaviors (e.g., modeling desired behavior, class discussion).	0.248	**0.659**	-0.096	0.008	-0.053
BAE	11. Academic expectations (e.g., meeting deadlines, completing assignments, studying for exams) set in this class are developmentally appropriate.	0.079	**0.655**	0.085	-0.132	-0.050
BAE	27. Students in this class understand what behavior is expected of them.	-0.042	**0.600**	-0.135	0.082	0.090
ORG	50. Students have access to the materials they need.	-0.009	**0.584**	-0.015	-0.094	0.194

Table 3.1 Continued

BAE	29. Students receive corrective feedback for inappropriate behaviors in a timely manner.	0.122	**0.494**	-0.096	0.055	0.196
ORG	54. Students are able to easily see the teacher and presented material during a class lesson.	-0.006	**0.448**	0.160	0.026	0.040
ORG	49. Students know where to find additional resources or supplies in the classroom (e.g., dictionaries, textbooks, pencils, paper).	0.029	**0.407**	0.101	0.022	0.212
ORG	59. Students are assigned a manageable amount of work each class.	0.109	**0.398**	0.081	-0.101	0.185

Factor 3 – Assessment and Instructional Learning

Table 3.1 Continued

AS	95. Students are guided in creating study guides to help them prepare for exams.	-0.034	-0.162	**0.707**	-0.041	0.034
AS	111. Students know how to improve their grades from the written feedback they receive.	0.039	-0.053	**0.690**	0.016	0.014
AS	96. Students are shown or taught how to study for tests and quizzes.	0.015	-0.014	**0.674**	-0.112	0.058
AS	100. Students are provided with learning objectives to ensure that they know what material is important.	0.078	-0.013	**0.608**	0.150	0.024
AS	116. Directions of how students can improve their grades are regularly communicated to caregivers.	0.232	0.019	**0.596**	0.019	-0.143

Table 3.1 Continued

AS	115. Students understand how assignments will be graded (e.g., scoring rubrics) prior to turning them in.	-0.099	0.167	**0.587**	-0.035	0.083
IT	118. Projects that require students to present course content in novel ways are regularly assigned.	0.226	-0.131	**0.536**	-0.179	-0.043
AS	97. Students are aware of what material they will be tested on.	-0.071	0.133	**0.500**	0.215	0.058
AS	114. Grading criteria and objectives for all assignments are clearly communicated to students.	-0.073	0.335	**0.436**	0.050	0.085

Table 3.1 Continued

AS	94. Students are given all the information they need to know in order to do well in this class (e.g. know when tests are scheduled, what material will be on tests, how to prepare for exams).	-0.138	0.190	**0.412**	0.133	0.198
IT	121. Students are taught how to summarize the key concepts of book chapters or lessons.	0.232	-0.066	**0.380**	-0.048	0.193
AS	103. Students' knowledge of class content is evaluated on a consistent and timely basis.	-0.122	0.244	**0.364**	-0.046	0.263

Table 3.1 Continued

IT	119. Students are required to present material to their peers in a way that encourages active problem solving or critical thinking.	0.294	0.063	**0.360**	-0.035	0.077
AS	106. Test content directly reflects learning objectives.	-0.172	0.306	**0.356**	-0.054	0.265
	Factor 4 - Classroom Chaos					
IT	130. It is difficult to find activities that keep students engaged. ®	0.014	-0.192	-0.018	**0.745**	0.158
IT	129. It is difficult to think of real-world examples that would resonate with students. ®	-0.026	-0.133	-0.054	**0.745**	0.166

Table 3.1 Continued

AS	113. It is not possible to provide students with timely feedback on assignments (e.g., homework, tests). ®	0.038	0.002	0.052	**0.705**	-0.159
BAE	38. Students in this class receive more negative feedback than positive feedback about their behavior. ®	0.050	-0.009	-0.003	**0.700**	-0.033
BAE	35. Students are unsure what will happen if they do not follow the rules. ®	-0.071	0.135	0.054	**0.595**	-0.097
BAE	6. It is unclear what students are expected to do when they come to class. ®	0.051	0.108	0.084	**0.577**	-0.242

Table 3.1 Continued

ORG	66. Classroom routines and procedures are dependent on teacher-related factors (e.g., teacher energy, preparedness). ®	-0.078	0.084	-0.026	**0.493**	-0.108
ORG	71. The classroom routine frequently requires adjustment. ®	-0.031	0.082	-0.157	**0.451**	0.052
	Factor 5 - Classroom Climate and Instructional Support					
IT	139. Students' responses are repeated or expanded upon, so they feel heard and participation is encouraged.	0.007	0.092	0.004	-0.168	**0.657**
IT	125. Questions are used as a way to stimulate class discussions.	-0.011	-0.001	0.083	-0.024	**0.643**

Table 3.1 Continued

IT	137. Previously learned content is worked into new lessons to provide repetition or to make connections across lessons.	-0.109	0.054	0.099	-0.055	**0.643**
IT	128. The use of real-world examples is used to make the material more meaningful or relevant.	-0.002	0.012	0.169	0.112	**0.553**
IT	127. A variety of methods (e.g., diagrams, videos, discussions, debates) are used to maintain student interest and to encourage participation.	0.052	0.048	0.121	0.035	**0.489**

Table 3.1 Continued

ESS	86. A warm and supportive environment is maintained for students in this class.	0.183	0.150	-0.098	0.178	**0.449**
ESS	87. When students demonstrate prosocial behaviors (e.g., explaining an assignment to a peer), they are acknowledged or praised.	0.146	0.049	0.021	0.020	**0.425**
IT	134. Students who need extra assistance are often paired with students who do well in the class when given group assignments.	0.223	-0.066	0.009	-0.165	**0.415**

Table 3.1 Continued

ESS	89. Students are encouraged to use appropriate problem-solving skills during conflicts with classmates.	0.280	0.217	-0.041	0.083	**<u>0.318</u>**
ESS	80. Students are treated with respect at all times.	0.110	0.186	0.089	0.066	**<u>0.282</u>**

Note. Text that is bolded and underlined indicates a primary component loading. ® indicates items that were reverse coded. The originally hypothesized factor names were: ESS = Emotional and Social Support; BAE = Behavioral and Academic Expectations; IT = Instructional Support; AS = Assessment; ORG = Organization.

Internal Reliability

Internal consistency was examined using Cronbach's alpha, as this is the most used and accepted reliability coefficient (Cronbach, 1951). Evidence of internal consistency is important in measure development because it provides information about whether the items are measuring the same underlying construct. Alpha values greater than .9 indicate excellent internal consistency, whereas alpha values that are at least .7 indicate acceptable levels of internal consistency (Blunch, 2008). Cronbach's alpha for the total score of the ATEP was .943. The internal consistencies ranged from .860 to .904 for the five factors of the ATEP, which provides good evidence of internal consistency. Specifically, the Cronbach's alpha was .904 for Emotional and Social Support (Factor 1), .897 for Classroom Expectations and Routines (Factor 2), .900 for Assessment and

Instructional Learning (Factor 3), .833 for Classroom Chaos (Factor 4), and .860 for Classroom Climate and Instructional Support (Factor 5).

In order to examine the relationship between the five factors of the ATEP, bivariate correlations for the ATEP domains were conducted (Table 3.2). According to Thomas (2014), there is an expectation that a measure's subscales would be correlated, as this suggests that subscales are generally assessing related constructs. However, caution is advised because correlations that are too high might indicate too much overlap between subscales, which would suggest that separate constructs are not being measured (Thomas, 2014). Emotional and Social Support (Factor 1) was found to be significantly and moderately correlated with Classroom Expectations and Routines (Factor 2; r = .48), Assessment and Instructional Learning (Factor 3; r = .54), Classroom Climate and Instructional Support (Factor 5, r = .61), and the ATEP total score (r = .682). Classroom Expectations and Routines (Factor 2) was moderately to strongly associated with Assessment and Instructional Learning (Factor 3; r = .68), Classroom Chaos (Factor 4; r = .359), Classroom Climate and Instructional Support (Factor 5; r = .69), and the ATEP total score (r = .85). The correlation between Assessment and Instructional Learning (Factor 3) and Classroom Chaos (Factor 4) was significant and small in magnitude (r = .16), but Assessment and Instructional Learning (Factor 3) had strong correlations with both Classroom Climate and Instructional Support (Factor 5; r = .71), and the ATEP total score (r = .83). Classroom Chaos (Factor 4) was moderately to strongly correlated with Classroom Climate and Instructional Support (Factor 5; r = .50), and the ATEP total

score (r = .83). Lastly, Classroom Climate and Instructional Support (Factor 5) was found to be strongly related to the ATEP total score (r = .85).

Table 3.2 *Bivariate correlations between ATEP scales.*

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	ATEP Total
Factor 1	--	--	--	--	--	--
Factor 2	.476***	--	--	--	--	--
Factor 3	.540***	.676***	--	--	--	--
Factor 4	-0.048	.359***	.158**	--	--	--
Factor 5	.612***	.698***	.713***	.241***	--	--
ATEP Total	.683***	.852***	.827***	.501***	.851***	--

Note: *p <.05, **p<.01, ***p<.001. Factor 1 - Emotional and Social Support; Factor 2 - Classroom Expectations and Routines; Factor 3 - Assessment and Instructional Learning; Factor 4 - Classroom Chaos; Factor 5 -Classroom Climate and Instructional Support.

Convergent Validity

The My Class Inventory – Short Form (TCMI-SF) was also used to provide evidence of convergent validity as it has a great deal of overlap with the ATEP in that it also assesses the classroom environment as it relates to students' relationships with one another and students' perceptions of difficulty of class material. It was expected that scores on the ATEP would be significantly and positively associated with TCMI-SF scores.

When evaluating whether there is evidence of convergent validity, it has been suggested that correlations above .70 provide evidence of convergent validity, whereas scores below .50 do not provide evidence of convergent validity (Field, 2015; Meyers et al., 2013). Results of correlations between the ATEP and TCMI-SF revealed moderately

strong associations between the TCMI-SF Satisfaction subscale and Emotional and Social Support (Factor 1; r = .525), Classroom Climate and Instructional Support (Factor 5; r = .543) and the Total score (r = .541) of the ATEP. Interestingly, a strong and negative correlation was found between Classroom Chaos (Factor 4) of the ATEP and the TCMI-SF Difficulty subscale (r = -.758).

Concurrent Validity

Past studies suggest that teachers who experience high levels of self-efficacy and do not feel burdened by their jobs are more likely to implement high quality teacher practices (Brouwers & Tomic, 2000; Cadavid & Lunenburg, 1991). Given these findings, concurrent validity was assessed by examining the extent to which the ATEP total score and factor scores correlated with measures designed to assess: (a) teachers' beliefs in their own abilities to perform teaching tasks (Teachers' Efficacy Beliefs System - TEBS) and (b) teachers' feelings of burnout (Maslach Burnout Inventory – Educators Survey; MBI-ES). It was expected that scores on the ATEP would be significantly and positively associated with scores on the TEBS whereas scores on the ATEP would be significantly and negatively associated with scores on the MBI-ES.

The Classroom Expectations and Routines (Factor 2) subscale of the ATEP was significantly and positively correlated with the following TEBS-S subscales: Communication/Clarification (r = .569), Management/Climate (r = .581), Individual Differences (r = .505), Students Motivation (r = .533), and TEBS-S Total score (r = .585). The Assessment and Instructional Learning (Factor 3) and Classroom Climate and Instructional Support (Factor 5) subscales of the ATEP had moderately strong correlations with the TEBS-S Total score (r = .501). Finally, the Total score of the ATEP

was moderately associated with the following TEBS-S subscales: Communication/Clarification (r = .571), Management/Climate (r = .560), Individual Differences (r = .502), Students Motivation (r = .538), Higher Order Thinking Skills (r = .516), and TEBS-S Total score (r = .588). Regarding bivariate correlations between scores on the ATEP and the MBI-ES, results revealed a significant and negative relationship of moderate strength between Classroom Chaos (Factor 4) of the ATEP and the MBI Depersonalization scale (r = -.609). The remaining significant relationships between the ATEP and MBI were below the .5 cut-off. Tables 4, 5, and 6 present bivariate correlations between all subscale and total scores of the ATEP, TCMI-SF, TEBS, and MBI.

Concurrent Validity

Past studies suggest that teachers who experience high levels of self-efficacy and do not feel burdened by their jobs are more likely to implement high quality teacher practices (Brouwers & Tomic, 2000; Cadavid & Lunenburg, 1991). Given these findings, concurrent validity was assessed by examining the extent to which the ATEP total score and factor scores correlated with measures designed to assess: (a) teachers' beliefs in their own abilities to perform teaching tasks (Teachers' Efficacy Beliefs System - TEBS) and (b) teachers' feelings of burnout (Maslach Burnout Inventory – Educators Survey; MBI-ES). It was expected that scores on the ATEP would be significantly and positively associated with scores on the TEBS whereas scores on the ATEP would be significantly and negatively associated with scores on the MBI-ES.

The Classroom Expectations and Routines (Factor 2) subscale of the ATEP was significantly and positively correlated with the following TEBS-S subscales:

Communication/Clarification (*r* = .569), Management/Climate (*r* = .581), Individual Differences (*r* = .505), Students Motivation (*r* = .533), and TEBS-S Total score (*r* = .585). The Assessment and Instructional Learning (Factor 3) and Classroom Climate and Instructional Support (Factor 5) subscales of the ATEP had moderately strong correlations with the TEBS-S Total score (r = .501). Finally, the Total score of the ATEP was moderately associated with the following TEBS-S subscales: Communication/Clarification (*r* = .571), Management/Climate (*r* = .560), Individual Differences (*r* = .502), Students Motivation (*r* = .538), Higher Order Thinking Skills (*r* = .516), and TEBS-S Total score (*r* = .588). Regarding bivariate correlations between scores on the ATEP and the MBI-ES, results revealed a significant and negative relationship of moderate strength between Classroom Chaos (Factor 4) of the ATEP and the MBI Depersonalization scale (*r* = -.609). The remaining significant relationships between the ATEP and MBI were below the .5 cut-off. Tables 3.3, 3.4, and 3.5 present bivariate correlations between all subscale and total scores of the ATEP, TCMI-SF, TEBS, and MBI.

Table 3.3 *Bivariate correlations examining convergent validity between the ATEP and the My Class Inventory subscales.*

	ATEP Factor 1	ATEP Factor 2	ATEP Factor 3	ATEP Factor 4	ATEP Factor 5	ATEP Total
TCMI-SF						
Satisfaction	**.525*****	.459***	.462***	0.070	**.543*****	**.541*****

Table 3.3 Continued

TCMI-SF Peer Relations	.329***	.229***	.250***	-0.025	.273***	.276***
TCMI-SF Competitiveness	.179**	.173**	.195***	-.243***	.166**	0.108
TCMI-SF Difficulty	0.079	-.297***	-0.097	**-.758*** **	-.198***	-.369***

Note: *p <.05, **p<.01, ***p<.001. Factor 1 - Emotional and Social Support; Factor 2 - Classroom Expectations and Routines; Factor 3 - Assessment and Instructional Learning; Factor 4 - Classroom Chaos; Factor 5 -Classroom Climate and Instructional Support.

Table 3.4 *Bivariate correlations examining convergent validity between the ATEP and the TEBS-S.*

	ATEP Factor 1	ATEP Factor 2	ATEP Factor 3	ATEP Factor 4	ATEP Factor 5	ATEP Total
TEBS-S Communication/ Clarification	.339***	**.569*** **	.468***	.275***	.485***	**.571*** **
TEBS-S Management/ Climate	.327***	**.581*** **	.452***	.258***	.485***	**.560*** **
TEBS-S Individual Differences	.370***	**.505*** **	.464***	.144*	.397***	**.502*** **

Table 3.4 Continued.

TEBS-S Student Motivation	.392**	**.533***	.435**	.204**	.456**	**.538***
TEBS-S HOTS	.374***	.485***	.471**	.136*	.486**	**.516***
TEBS-S Total	.397***	.585***	.501**	.215**	.512**	**.588***

Note: *p <.05, **p<.01, ***p<.001. Factor 1 - Emotional and Social Support; Factor 2 - Classroom Expectations and Routines; Factor 3 - Assessment and Instructional Learning; Factor 4 - Classroom Chaos; Factor 5 -Classroom Climate and Instructional Support.

Table 3.5 *Bivariate correlations examining convergent validity between the ATEP and the Maslach Burnout Inventory – Educator*

	ATEP Factor 1	ATEP Factor 2	ATEP Factor 3	ATEP Factor 4	ATEP Factor 5	ATEP Total
MBI Emotional Exhaustion	-0.087	-.123*	-0.112	-.306***	-0.105	-.212***
MBI Deperson-alization	-0.033	-.197**	-0.084	**-.609*** **	-.188**	-.322***

Table 3.5 Continued.

MBI Personal Accomplish-ment	.484***	.354***	.423***	0.039	.434***	.459***
MBI Total	0.097	-0.051	0.035	-.462***	-0.008	-.127*

Note: *p <.05, **p<.01, ***p<.001. Factor 1 - Emotional and Social Support; Factor 2 - Classroom Expectations and Routines; Factor 3 - Assessment and Instructional Learning; Factor 4 - Classroom Chaos; Factor 5 -Classroom Climate and Instructional Support.

Discriminant Validity

Finally, discriminant validity was assessed by examining the extent to which scores on the ATEP correlated with scores on the Marlowe-Crowne Social Desirability Scale (MCSDS; see Table 3.6), as these measures were thought to assess different and unrelated constructs. Results revealed significant and positive correlations that were small in magnitude between the total score of the MCSDS and four of the factor scores and total score of the ATEP (*r's* =.162-.235).

Table 3.6 *Bivariate correlations examining discriminant validity between the ATEP and the Marlowe-Crowne Social Desirability Scale (MCDDS).*

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	ATEP Total
MCDDS	0.080	.169**	.195**	.235***	.162**	.235***

Note: *p <.05, **p<.01, ***p<.001. Factor 1 - Emotional and Social Support; Factor 2 - Classroom Expectations and Routines; Factor 3 - Assessment and Instructional Learning; Factor 4 - Classroom Chaos; Factor 5 -Classroom Climate and Instructional Support.

CHAPTER IV – DISCUSSION

The extant literature offers strong support for the positive impact that teacher practices have on students' academic achievement and classroom behaviors (Andersson & Palm, 2016; Oliver, Wehby, & Reschly, 2011; Rimm-Kaufman, Storm, Sawyer, Pianta, & LaParom, 2006; Rubie-Davies, 2007; Sandholtz, 2011). Therefore, the importance of teacher practices cannot be understated, but rather comprehensively and reliably evaluated so that teachers can reflect on how their own practices may be improved. To date, the only tools available to comprehensively assess teacher practices are observational systems, which require a great deal of resources (i.e., time and money) and thus are often not viable for school districts to administer. Existing teacher-report measures also present concerns, as these tools often do not fully assess the multi-faceted nature of teacher practices or have poor psychometric properties. To address these concerns, the present study added to the extant literature by developing the Assessment of Teacher Practices and Assessment (ATEP) measure to assess teacher practices feasibly, comprehensively, reliably, and validly. Furthermore, the ATEP is the first known measure developed with the sole intention of offering teachers a self-evaluative method to assess their teaching practices so they may identify areas of further growth and track their progress over time.

According to Goe's (2007) Framework of Teacher Quality, effective teacher practices are a vital and modifiable component comprising teacher quality. Teacher practices are arguably the most important component of teacher quality, when considering the impact that they have on student behaviors and achievement and the fact that they can be easily targeted through teacher development trainings. Teacher practices

include teacher-specific behaviors inclusive of classroom management, organization, instructional delivery, and expectations. A main objective in the development of the ATEP was to comprehensively assess teacher practices that have been linked to students' academic and behavioral success. To achieve this objective, a thorough review of the research literature and other existing measures of teacher practices (i.e., behavioral observations and teacher/student report) was conducted to ensure all facets of this construct was assessed by the ATEP. Following this review, five domains of teacher practices stood out and were consistently identified as important for student success. These domains represented the extent to which teachers (a) communicate behavioral and academic expectations to their students (Behavioral and Academic Expectations), (b) maintain a consistent classroom structure, pace, and routine (Organization), (c) encourage emotional development of their students and foster a positive classroom environment (Emotional and Social Support), (d) assess their students learning of course content and prepare them for subsequent evaluation (Assessment), and (e) encourage higher order thinking skills, engagement, and the development of skills helpful for retaining new material/knowledge (Instructional Techniques).

Content validity of the ATEP was evaluated by three experts in the field of education. As expected, all three experts rated the originally proposed domains of Behavioral and Academic Expectations, Emotional and Social Support, Assessment, and Instructional Techniques as essential and capturing the construct of teacher practices. In contrast, only one expert rated the Organization domain as essential. Regarding individual items, items on the Behavioral and Academic Expectations domain that were rated as unessential focused on narrowly defined academic expectations (due dates, high

quality versus low quality work, and format/neatness of assignments) that may not be deemed important by most teachers and very specific forms of consequences or feedback that are given in response to student disruptive behaviors. Items on the Organization domain that were rated as unessential tended to capture the pace of classroom instruction or if the workload was appropriately balanced with the time allocated to complete it. Lastly, items rated as unessential from the other domains tended to be redundant and overlapped with other items that better captured the construct of interest. In general, items rated as unessential might have been too specific and narrow (e.g., types of reinforcement) or might be perceived as the bare minimum in terms of what teachers should already be doing. Experts found the remaining items to be well-written and comprehensive given that they did not suggest changes to the existing items, nor did they suggest additional items be added.

In alignment with the number of domains that were theorized to comprise the ATEP, the results of an Exploratory Factor Analysis also revealed a five-factor structure of the ATEP albeit the items loaded somewhat differently than expected. Factor 1 was comprised almost entirely of items that were developed to capture the Emotional and Social Support domain, with the exception of one item from the Behavioral and Academic Expectations (BAE) domain that also loaded onto this factor. The items that loaded onto Factor 2 comprised mostly of items from the BAE domain, but also contained four items from the Organization domain. To reflect this updated composition of items, Factor 2 was named Classroom Expectations and Routines, as it assessed the extent to which teachers communicated behavioral and academic expectations to students and made efforts to provide a stable and predictable classroom routine. Factor 3

comprised of items from two domains: Assessment and Instructional Techniques. As a result, it was labeled Assessment and Instructional Learning considering these items captured the extent to which teachers assessed their students learning and helped students develop effective study and critical thinking skills. Unexpectedly, the items that loaded onto Factor 4 were all the items that were reverse coded on the original ATEP and came from four domains (i.e., Behavioral and Academic Expectations, Organization, Assessment, and Instructional Techniques). This factor was named Classroom Chaos, as it assessed teacher's difficulty in maintaining a classroom environment that is conducive to learning. Lastly, Factor 5 comprised of items from the Instructional Techniques and Emotional and Social Support domains. It was named Classroom Climate and Instructional Support, as it captured teacher's ability to maintain an engaging and supportive classroom environment.

Results from this study also provided good evidence of internal consistency, and some evidence of concurrent, convergent, and discriminant validity. Cronbach's alpha for the ATEP subscales ranged from .86 to .90 and are comparable or outperform existing measures assessing teacher practices. The associations between the ATEP's factors were moderate to strong in strength indicating they are measuring the same construct while still contributing unique information. There was one exception in that Factor 4 (Classroom Chaos) displayed nonsignificant to modest correlations ranging from -.048 (Emotional and Social Support) to .241 (Classroom Climate and Instructional Support). However, Factor 4 (Classroom Chaos) was moderately correlated with Factor 2 (Classroom Expectations and Procedures). This pattern of results makes sense given the

item content of (Classroom Chaos) Factor 4 reflects an absence of routines and engagement in effective instructional practices.

Regarding evidence of convergent validity, correlations between subscales on the My Class Inventory – Short Form for Teachers (TCMI-SF) and ATEP factors ranged from weak to strong; however, not all subscales of the TCMI-SF appear to align with the domains of the ATEP. For example, the subscales with nonsignificant to weak correlations with the ATEP assess the quality of student relationships including their level of competitiveness (TCMI-SF Peer Relations and Competitiveness). In contrast, TCMI-SF subscales with moderate to strong correlations have item content that do overlap with the ATEP factors; specifically, the Emotional and Social Support (Factor 1) and Classroom Climate and Instructional Support (Factor 5) of the ATEP were moderately associated with the TCMI-SF Satisfaction subscale. Factor 4 of the ATEP (Classroom Chaos) was strongly associated with the TCMI-SF Difficulty subscale. In examining discriminant validity, very modest correlations (r's < .25) were found between the total score of the Marlowe-Crowne Social Desirability Scale (MCSDS) and the total score and subscale scores of the ATEP. In sum, results provided evidence of convergent and discriminant validity for the ATEP.

Concurrent validity was examined between the ATEP and the Teachers' Efficacy Beliefs System (TEBS-S) and the Maslach Burnout Inventory- Educators Survey (MBI-ES). As expected, results showed a moderately strong relationship between the TEBS-S Total score and the ATEP Total score providing evidence of concurrent validity. Although a weak relationship was found between the total score of the Maslach Burn out Inventory – Educators Survey (MBI-ES) and the total score of the ATEP, a significant

and negative correlation of moderate strength was found between Classroom Chaos (Factor 4) of the ATEP and the MBI-ES Depersonalization subscale. This pattern of results makes sense considering that the MBI-ES Depersonalization subscale has traditionally reflected the attitudes often associated with burnout (Maslach et al., 2001) and is classified as "mental distance" from students and/or the profession of teaching (Simbula & Guglielmi, 2010). These difficulties would certainly contribute to difficulty engaging in effective teacher practices and maintaining expectations and routines. Results provide evidence of concurrent validity, thus indicating that the ATEP is a promising teacher-report measure that should continue to be evaluated.

Limitations

Despite the present study's promising results, some limitations should be considered. First, the sample size of the present study was somewhat small when considering what some experts have recommended in the extant literature. For example, Gorsuch (1983) suggests that there should be five to ten participants per questionnaire item when conducting an EFA, so this study's sample size should have fallen in the range of 695 to 1,390 participants. However, others have argued that using an absolute or ratio-based guideline is a flawed approach, as it assumes that EFAs conducted with smaller sample sizes produce inaccurate results (Arrindell & van der Ende, 1985; Jackson, 2001; MacCallum, Widaman, Zhang, & Hong, 1999). Indeed, MacCallum and colleagues (1999) performed a series of Monte Carlo simulations using artificial data and determined that sample sizes ranging from 150 to 250 participants are sufficient in garnering accurate and unbiased factor loadings under certain conditions (e.g., when communalities are at or above the .5 range; MacCallum, Widaman, Zhang, & Hong, 1999). Thus, future studies

should consider the various recommendations in the literature to ensure an appropriate sample size.

Furthermore, limitations exist with the measures included for convergent and concurrent validity (My Class Inventory – Short Form for Teachers, TCMI-SF; Maslach Burnout Inventory – Educator Survey, MBI-ES). The TCMI-SF, although originally thought to assess similar domains as the ATEP, upon closer examination only had two subscales that overlapped with the ATEP (Satisfaction and Difficulty). The TCMI-SF was rather narrow in its scope with respect to item content, a limitation previously noted of existing teacher-rated measures, and primarily focused on teacher's perceptions of students' relationships in the classroom. Therefore, future studies should consider making use of several self-report teacher measures that align with all the domains of the ATEP to obtain evidence of convergent validity. Lastly, although the MBI-ES displayed desirable results ($r > .50$) supporting concurrent validity between Classroom Chaos (Factor 4) of the ATEP and MBI-ES Depersonalization subscales, future studies should consider including other measures to establish concurrent validity.

An additional limitation to consider is that data was collected during the fall semester of 2020, which was when there were differences pertaining to whether classes were in-person, virtual, or through a hybrid combination (in-person and virtual). For the present sample, 37.2% (N = 100) reported teaching only virtually, 39.8% (N = 107) taught in a hybrid format (in-person and virtual), and 23% (N = 62) taught fully in-person. As such, teachers completing our study survey were instructed to formulate their responses to reflect their general in-person practices. Literature suggests that recall bias may result in inaccurate data, as information might be omitted or simply not remembered

(Coughlin, 1990; Talari & Goyal, 2020). Participants who were teaching fully remotely/virtually were asked to respond retrospectively and thus there is a chance participants misremembered or omitted information. It cannot be ruled out that retrospective reporting of teacher practices impacted results.

The present sample was representative of the U.S. population of teachers regarding two demographic characteristics (ethnicity and gender). However, efforts should always be made to recruit samples that are representative of the population as this will ensure that measures are generalizable to teachers of diverse backgrounds. For the present sample, participants were largely homogenous in terms of racial identify. Specifically, 84% (N = 226) of participants identified as White, whereas data from the Institute of Education Sciences reported 79% of U.S. teachers identify as White.

Future Directions

Before the ATEP can be used as a progress measure to track improvements in teacher practices, other measure development studies should be conducted. First, future studies should focus on verifying the structure of the ATEP proposed by the EFA through a confirmatory factor analysis (CFA) in a new sample. Additionally, psychometric properties of the ATEP should also be re-evaluated in a sample that is representative of teachers in the U.S. Specifically, examining convergent validity between the ATEP and a comprehensive observational measure of teacher practices (e.g., CLASS) would be vital in establishing whether the ATEP could assess the multiple facets of teacher practices as well as observational measures that require a good deal of resources to reliably administer.

Other psychometric properties of interest that should be evaluated in future studies include test-retest reliability and criterion validity (i.e., concurrent and predictive). Indeed, evidence of test-retest reliability is vital if the ATEP is to be used as a progress monitoring measure. Future studies should use caution in examining test-retest as it is important for the ATEP to also be sensitive to intervention effects; therefore, test-retest should be conducted within a short interval (i.e., one to two months) prior to interventions targeting teacher practices. Additionally, to examine whether the ATEP is sensitive to treatment effects, researchers might consider examining changes in teacher practices for teachers receiving an intervention to improve practices and those not receiving an intervention. Concurrent validity was examined in the present study by conducting correlations between scores on the ATEP and measures of teacher's self-efficacy and burnout, as these constructs were hypothesized to be related to teacher practices. For example, it might be useful to examine the relationship between the ATEP and students' classroom behavior infractions since there might be a relationship between poor teacher practices and high rates of disruptive classroom behaviors. Predictive validity is equally as important because it would indicate whether our measure of teacher practices can predict student outcomes as well as other measures of teacher practices. Future studies could then examine what domains of teacher practices are most predictive of student success, which would be vital in informing teacher trainings, as teacher practices could be targeted that offer the most benefit for students.

Conclusion

The primary goal of the present study was to develop a comprehensive measure that reliably and validly assessed teacher practices and expectations. A thorough literature

review, along with expert feedback, resulted in a set of items that thoroughly covered all aspects of teacher practices and expectations. Exploratory factor analyses supported a five-factor structure for the final 58 items of the Assessment of Teacher Expectations and Practices, with high factor loadings and estimates of internal consistencies for each factor provided strong evidence of reliability. Additionally, results of bivariate correlations provided moderate evidence of convergent validity and discriminant validity. Future studies should verify the structure of the ATEP and re-examine convergent and concurrent validity with different measures. Despite the need for further research, these findings are promising in that the proposed measure may be a time and cost-efficient alternative to existing observational measures to monitor the use of teacher practices in the classroom that are expected to promote the academic and behavioral success of students.

APPENDIX A – IRB

Date: 6-15-2021

IRB #: IRB-19-587
Title: ATEP Measure Development
Creation Date: 12-3-2019
End Date:
Status: Approved
Principal Investigator: Kimberly Barajas
Review Board: Sacco (Exempt/Expedited Board)
Sponsor:

Study History

Submission Type Initial	**Review Type** Expedited	**Decision** Approved
Submission Type Modification	**Review Type** Expedited	**Decision** Approved
Submission Type Modification	**Review Type** Expedited	**Decision** Approved

Key Study Contacts

Member Stephanie Smith	**Role** Co-Principal Investigator	**Contact** sd.smith@usm.edu
Member Kimberly Barajas	**Role** Primary Contact	**Contact** kimberly.barajas@usm.edu
Member Kimberly Barajas	**Role** Principal Investigator	**Contact** kimberly.barajas@usm.edu

APPENDIX B - My Class Inventory – Short Form for Teachers

1. The students enjoy their schoolwork in the class.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

2. Students do not fight with each other.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

3. Students often race to see who can finish their work first.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

4. In the class the work is hard to complete.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

5. In the class everyone is friends.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

6. The school counselor helps students feel good about learning in this classroom.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

7. Students are happy with the class.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

8. Most students want their work to be better than their friend's work.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

9. Most students cannot complete their assignments without a lot of help.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

10. The school counselor aids with building classroom cohesion.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

11. Students in the class have good buddies.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

12. Students seem to like the class.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

13. Only the brightest students can do all the work.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

14. Because of the school counselor's visits to the classroom, the students tend to work more cooperatively.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

15. All students in my class get along well with each other.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

16. Most students appreciate their learning experiences in the class.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

17. Some students always try to outperform their peers.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

18. The schoolwork is too complicated for the students.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

19. The school counselor helps make the learning less difficult.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

20. All students in the class are fond of one another.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

21. The students see the class as fun.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

22. Students in the class do not argue with each other.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

23. Most students in the class do not know how to do their work very well.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

24. The school counselor helps create unity in the classroom.

1	2	3	4	5
Strongly Disagree		Neutral		Strongly Agree

APPENDIX C - Maslach Burnout Inventory – Educators Survey

1. I feel frustrated by my job.

_____

1 2 3 4 5 6 7

Never Always

2. Working with people all day is really a strain for me.

_____

1 2 3 4 5 6 7

Never Always

3. Working with people directly puts too much stress on me.

_____

1 2 3 4 5 6 7

Never Always

4. I feel burned out from my work.

_____

1 2 3 4 5 6 7

Never Always

5. I feel emotionally drained from my work.

_____

1 2 3 4 5 6 7

Never Always

6. I feel like I'm at the end of my rope.

_____

1 2 3 4 5 6 7

Never Always

7. I feel fatigued when I get up in the morning and have to face another day on the job.

_____

1 2 3 4 5 6 7

Never Always

8. I feel I'm working too hard on my job.

_____

1 2 3 4 5 6 7

Never Always

9. I feel used up at the end of the workday.

_____

1 2 3 4 5 6 7

Never Always

APPENDIX D - Teachers' Efficacy Beliefs System - Self

Directions: Read each question carefully. Consider each response from your own teaching experiences. Please indicate your frank opinion by circling the appropriate number. Your responses will remain confidential.

1. How much can you express your views freely on important school matters?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

2. How much do you feel you can influence the learning of underachieving students?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

3. How much can you do to get students to follow classroom rules?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

4. How much do you connect learning to students' cultural backgrounds?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

5. How much can you do to make the school a safe place?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

6. How much can you do to get students to believe they can do well in schoolwork?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

7. How much can you do to get parents to become involved in their children's learning?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

8. How much can you influence the decisions that are made in the school?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

9. How much do you help students to question ideas based on their own cultural backgrounds?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

Teacher Efficacy Survey - continued

10. How much does student disciplinary issues interfere with instruction?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

11. How much do you know and understand the cultural practices and beliefs that are in your immediate school community?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

12. How much can you help other teachers with their teaching skills?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

13. How much do you have your classroom environment reflect students' cultures?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

14. How much can you do to reduce school absenteeism?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

15. How much can you do to make students enjoy coming to school?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

16. How much can you do to enhance collaboration between teachers and the administration to make the school run effectively?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

17. How much can you do to promote learning when there is a lack of support from the home?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

18. How much can you do to get through to the students with the most disciplinary issues?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

19. How much can you do to have your curricular materials reflect your students' cultural backgrounds?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

Teacher Efficacy Survey - continued

20. How much can you do to keep students on task on difficult assignments?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

21. How much do students' socioeconomic and/or cultural backgrounds influence your expectations of them?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

22. How much can you do to provide students with more individualized methods of instruction?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

23. How much can you do to effectively utilize instructional strategies that promote positive self-images in students?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

24. How much can you do to maximize all learning experiences for struggling students?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

25. How much do you understand the cultural and social forces that impede student learning?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

26. How much can you do to motivate students who show low interest in schoolwork?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

27. How much can you do to understand the attitudes and beliefs of your students and families that may impact student achievement?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

28. How much can you do to understand the historical and cultural aspects of your immediate school community?

1	2	3	4	5	6	7	8	9
Nothing		Very Little		Some Influence		Quite a Bit		A Great Deal

APPENDIX E - Marlowe-Crowne Social Desirability Scale

1. It is sometimes hard for me to go on with my work if I am not encouraged.
   a. True
   b. False
2. I sometimes feel resentful when I don't get my way.
   a. True
   b. False
3. On a few occasions, I have given up doing something because I thought too little of my ability.
   a. True
   b. False
4. There have been times when I felt like rebelling against people in authority even though I knew they were right.
   a. True
   b. False
5. No matter who I am talking to, I'm always a good listener.
   a. True
   b. False
6. There have been occasions when I took advantage of someone.
   a. True
   b. False
7. I am always willing to admit when I make a mistake.
   a. True
   b. False
8. When I don’t know something, I don’t at all mind admitting it.
   a. True
   b. False
9. I am always courteous, even to people who are disagreeable.
   a. True
   b. False
10. I have never been irked when people expressed ideas very different from my own.
    a. True
    b. False
11. There have been times when I was quite jealous of the good fortune of others.
    a. True
    b. False
12. I am sometimes irritated by people who ask favors of me.
    a. True
    b. False
13. I have never deliberately said something that hurt someone’s feelings.
    a. True
    b. False

APPENDIX F - Teacher Demographic Form

What is your age? 20-30 31-40 41-50 51-60 60+

What is your gender? Male Female

What is your race? White Black Asian More than one

What is your ethnicity? Hispanic Not Hispanic

What is your marital status?_____

Family income:_____

How many years have you been in the field of education? _____

How many years have you been a teacher? _____

What previous grades have you taught?_____

How many years have you worked in this school? _____

Do you teach at a private or public school (or other)? Please indicate here._____

Estimated number of students who receive free or reduced lunch: _____

Estimated number of total students in each grade:_____

Size of the current classroom:_____

Areas of specialization:_____

Degrees completed (Check all that apply)

____ Associate's degree

____ Bachelor's degree

____ Master's degree

____ Postmaster's certificate

____ Doctorate

Concentration or major: _____

What type of teacher are you? (Circle one) Regular/General Special Education

What grade(s) do you teach? (check all that apply)

Grade 1	Grade 5	Grade 9
Grade 2	Grade 6	Grade 10
Grade 3	Grade 7	Grade 11
Grade 4	Grade 8	Grade 12

Which of the following describes the teaching certificate you currently hold in this state? (check all that apply)

____ Regular or standard state certificate or advanced professional certificate

____ Probationary certificate (issued after satisfying all requirements except the completion of a probationary period)

____ Provisional or other type of certificate given to persons who are still participating in what the state calls an "alternative certification program"

____ Temporary certificate (requires some additional college coursework, student teaching, and/or passage of a test before regular certification can be obtained)

____ Waiver or emergency certificate (issued to persons with insufficient teacher preparation who must complete a regular certification program in order to continue teaching)

____ I do not have any of the above certifications in THIS state

In the past 12 months, did you participate in any of the following professional development activities (check all that apply)?

____ University course(s) related to teaching

____ Observational visits to other schools

____ Workshops, conferences or training sessions in which you were a presenter

____ Other workshops, conferences or training sessions in which you were NOT a presenter

Excluding the training you received for the purposes of this study, how many hours of training or professional development on classroom behavior management have you had in the past 12 months?

____ None

____ 1-2 hours

____ 3-5 hours

____ 6-7 hours

____ 8+ hours

What classroom management techniques have you used in your classroom in the past 12 months to increase prosocial behaviors and limit disruptive behaviors (e.g., Time to Teach, Token System, Good Behavior Game, Traffic Light, Check-in/Check-out)?

_____

APPENDIX G – Descriptive Statistics for Teachers and Students

	Teachers	Students
	M (SD)	M (SD)
Number of years teaching	14.57 (9.829)	--
Students in each grade	--	263.09 (936.28)
Students receiving free or reduced lunch		365 (939.02)
Race/Ethnicity	N (%)	Estimated Percentage
Black	20 (7.4%)	21.06%
White	226 (84%)	47.30%
Latinx/Hispanic	17 (6.3%)	20.46%
Asian	15 (5.65%)	7.96 %
Indigenous	2 (.7%)	3.33 %
Other	6 (2.2%)	--
School Type		
Public	221 (82.2%)	--
Private	40 (14.9%)	--
Charter	8 (3%)	--

APPENDIX H - Original Proposed Domain Names and Definitions

Behavioral and academic expectations: the teacher (a) clearly specifies the course content or learning objectives that will be covered and completed during class (items 1 – 8); (b) sets high work standards for all students; specifically, students know what work needs to be done, when it needs to be completed by (e.g., due dates), how to do the work well (e.g., high quality vs. low quality work; quality vs. quantity), and how to organize assignments (e.g., format, neatness) (items 9 – 20); (c) teacher describes and demonstrates desired behaviors (items 21 – 28); (d) teacher outlines clear consequences and provides feedback on student behaviors (items 29 – 46).

Organization: the teacher (a) maintains a well-organized classroom (e.g., class materials, physical structure; items 47 – 55); (b) facilitates an appropriate classroom pace (items 56 – 63); (c) sets and maintains classroom routines (items 64 – 71).

Emotional and Social Support: the teacher: (a) helps students to recognize, label, regulate, and express emotions (items 72 – 79); (b) provides a supportive student-teacher relationship and classroom environment (items 80 – 86), and (c) uses appropriate interpersonal effectiveness skills via positive feedback (items 87 – 93).

Assessment: the teacher (a) ensures that students know what material they will be tested on and how to prepare for exams (items 94 – 100); (b) monitors student learning by assessing what is retained (e.g., classroom assignments, tests, projects; items 101 – 110); (c) provides consistent, immediate, and specific feedback concerning students' academic performance (items 111 – 117).

Instructional Techniques: the teacher makes use of instructional techniques that (a) develop students' abilities to engage in higher-order thinking (e.g., critical thinking,

problem solving, reasoning skills; items 118 - 125), (b) increase the engagement of students (items 126 – 131), and (c) help students learn/retain new material (items 132 – 139).

APPENDIX I – Descriptive Statistics of ATEP Original Item Pool – 139 Items

Question Number	Item	CVR	Mean	Standard Deviation	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
BAE_1	Well-written and assessable learning objectives are clearly communicated to students.	0.333	3.4	0.545	0.469	0.956
BAE_2	Students have a clear understanding of what tasks need to be completed in class.	0.333	3.5	0.528	0.439	0.956
BAE_3	On any given day, students know what topics will be covered in class.	0.333	3.35	0.574	0.443	0.956
BAE_4	Predictability of what material will	-1	3.22	0.648	0.402	0.956

	be covered in class is of great importance.					
BAE_5	Students understand what they are expected to accomplish each class.	0.333	3.47	0.576	0.521	0.956
BAE_6	It is unclear what students are expected to do when they come to class. ®	0.333	3.25	0.896	0.182	0.957
BAE_7	In this class, expectations for completing academic work are different from day-to-day. ®	-0.333	2.69	0.885	0.017	0.957
BAE_8	Students have the necessary knowledge and skills	0.333	3.13	0.57	0.327	0.956

	to complete their work.					
BAE_9	Expectations for completing classwork are the same for all students.	-0.333	2.75	0.814	0.114	0.957
BAE_10	Instructions for class and homework assignments are clearly outlined for students.	1	3.53	0.527	0.63	0.956
BAE_11	Academic expectations (e.g., meeting deadlines, completing assignments, studying for exams) set in this class are developmentally appropriate.	1	3.45	0.526	0.588	0.956

BAE_12	Students are aware of how they should be performing academically (e.g., what grades are expected, subject mastery).	-1	3.34	0.61	0.507	0.956
BAE_13	Students know how to produce high-quality work (e.g., checking answers, proof-reading assignments, asking for clarification).	1	3.07	0.665	0.429	0.956
BAE_14	Students know in what format all their assignments should be done.	0.333	3.34	0.593	0.543	0.956
BAE_15	Students are aware tha work should be neat a easy to read.	-0.333	3.36	0.553	0.506	0.956

BAE_16	Due dates for tests, quizzes, and homework remain the same week-to-week.	-1	2.74	0.835	0.249	0.957
BAE_17	Students are regularly reminded about upcoming due dates.	-0.333	3.49	0.593	0.565	0.956
BAE_18	Students are expected to keep track of when they have tests or when assignments are due.	-0.333	3.23	0.659	0.335	0.956
BAE_19	In this class, the quality of the work produced by students is more important than the quantity.	0.333	3.4	0.637	0.421	0.956
BAE_20	The procedures for turning in class	0.333	3.53	0.577	0.527	0.956

	assignments or homework remain consistent.					
BAE_21	Classroom rules are clearly stated to students.	1	3.7	0.517	0.565	0.956
BAE_22	Students know what they are and are not allowed to do in this class.	0.333	3.66	0.485	0.518	0.956
BAE_23	Expectations for student behavior change on a regular basis. ®	0.333	3.13	0.982	0.098	0.957
BAE_24	Behavioral expectations for this class are developmentally appropriate (e.g., raise hand to speak, follow instructions).	1	3.58	0.572	0.543	0.956

BAE_25	It is often difficult for students to follow classroom rules. ®	1	2.86	0.833	0.174	0.957
BAE_26	Students are provided with examples of appropriate behaviors (e.g., modeling desired behavior, class discussion).	1	3.42	0.563	0.585	0.956
BAE_27	Students in this class understand what behavior is expected of them.	0.333	3.55	0.525	0.621	0.956
BAE_28	In general, behavioral expectations are the same for all students.	1	3.31	0.616	0.303	0.956

BAE_29	Students receive corrective feedback for inappropriate behaviors in a timely manner.	1	3.44	0.566	0.578	0.956
BAE_30	Students receive verbal praise when they follow classroom rules.	-0.333	3.4	0.63	0.502	0.956
BAE_31	Differential attention (e.g., ignoring inappropriate behaviors & providing attention to appropriate behaviors) is used to address minor behaviors.	0.333	3.16	0.617	0.399	0.956
BAE_32	Minor rule violations or disruptive behaviors are redirected by	-0.333	3.25	0.54	0.508	0.956

	focusing on behaviors that approximate the desired outcome.					
BAE_33	There are some negative behaviors that are never ignored (e.g., aggression) and have clear consequences.	1	3.6	0.607	0.379	0.956
BAE_34	Consequences for breaking classroom rules are consistent across all students.	-0.333	3.28	0.669	0.447	0.956
BAE_35	Students are unsure what will happen if they do not follow the rules. ®	1	3.26	0.878	0.29	0.956
BAE_36	Students receive different consequences for breaking minor rules	1	3.29	0.773	0.043	0.957

	(e.g., being off-task) versus breaking major rules (e.g., fighting with another student).					
BAE_37	Attention for appropriate behaviors is given more often than consequences for inappropriate behaviors.	1	3.07	0.655	0.422	0.956
BAE_38	Students in this class receive more negative feedback than positive feedback about their behavior. ®	0.333	3.1	0.884	0.239	0.957
BAE_39	It is difficult to provide praise or acknowledgement to all students for	-0.333	2.81	0.886	0.243	0.957

	appropriate behaviors. ®					
BAE_40	Students are noticed when they do something well.	1	3.43	0.565	0.57	0.956
BAE_41	Students have difficulty understanding what they need to do to receive positive feedback (e.g., praise, acknowledgement). ®	-0.333	3.05	0.837	0.313	0.956
BAE_42	At least one positive interaction occurs between each student and the teacher every class.	-0.333	3.08	0.72	0.457	0.956
BAE_43	It is easy to find positive student behaviors to praise.	0.333	3.21	0.716	0.403	0.956

BAE_44	Students are given tangible rewards (e.g., prizes, school supplies) for exceptionally good behavior.	-0.333	2.82	0.907	0.196	0.957
BAE_45	Students do not receive rewards for appropriate classroom behavior. ®	-0.333	2.68	0.927	0.186	0.957
BAE_46	Students feel motivated to make good classroom decisions.	1	3.25	0.57	0.51	0.956
ORG_47	Class materials are well-organized.	-0.333	3.4	0.584	0.593	0.956
ORG _48	Class materials are easily accessible.	-0.333	3.5	0.51	0.569	0.956
ORG _49	Students know where to find additional resources	0.333	3.44	0.566	0.589	0.956

	or supplies in the classroom (e.g., dictionaries, textbooks, pencils, paper).					
ORG _50	Students have access to the materials they need.	1	3.46	0.56	0.57	0.956
ORG _51	There are designated classroom areas for different class activities or specific resources (e.g., area for reading materials, math manipulatives or supplies).	-0.333	3.16	0.741	0.421	0.956
ORG _52	Students sit in the same seat every day.	-0.333	3.1	0.772	0.282	0.956
ORG _53	It is determined ahead of time where	0.333	3.12	0.727	0.245	0.956

	students sit in the classroom.					
ORG _54	Students are able to easily see the teacher and presented material during a class lesson.	1	3.56	0.549	0.575	0.956
ORG _55	Every aspect of this class is structured in a consistent and organized way.	0.333	3.25	0.638	0.582	0.956
ORG _56	Students have the same amount of work assigned to them each day.	-1	2.43	0.74	0.178	0.957
ORG _57	There are some days when students have more work than other days. ®	-1	1.92	0.591	-0.109	0.957
ORG _58	Students complete their work at the same time.	-1	2.24	0.891	0.033	0.957

ORG _59	Students are assigned a manageable amount of work each class.	-1	3.31	0.561	0.6	0.956
ORG _60	Students work at a similar pace on their in-class assignments.	-0.333	2.47	0.796	0.078	0.957
ORG _61	Students do not feel rushed to complete their work in class.	0.333	2.96	0.543	0.329	0.956
ORG _62	Students must take their work home because they are unable to complete it in class. ®	1	2.59	0.797	0.123	0.957
ORG _63	Students finish assignments at different times. ®	-0.333	1.75	0.586	-0.271	0.957
ORG _64	Class consistently starts and ends at the same time.	1	3.52	0.585	0.503	0.956

ORG _65	The daily classroom routine or agenda is posted so that students are able to reference it.	1	3.34	0.713	0.483	0.956
ORG _66	Classroom routines and procedures are dependent on teacher-related factors (e.g., teacher energy, preparedness). ®	0.333	2.67	0.913	0.186	0.957
ORG _67	Daily classroom routines and procedures stay the same.	-0.333	3.29	0.615	0.461	0.956
ORG _68	Class lessons are structured and ordered in the same way.	0.333	3.07	0.672	0.366	0.956

ORG _69	Classroom routines are consistent day-to-day.	0.333	3.27	0.58	0.597	0.956
ORG _70	Students complete similar types of assignments in the same order each day.	-0.333	2.84	0.759	0.289	0.956
ORG _71	The classroom routine frequently requires adjustment. ®	1	2.5	0.791	0.1	0.957
ESS_72	Students are taught how to label their emotions through reflection (e.g., “I can tell you are frustrated.”).	1	2.92	0.621	0.383	0.956
ESS_73	Students are encouraged to take the emotional perspective of others (e.g., “what do you	0.333	3.02	0.637	0.437	0.956

	think this person is feeling?”).					
ESS_74	Students feelings are normalized and validated (e.g., “Everyone gets angry.”).	0.333	3.18	0.568	0.448	0.956
ESS_75	Students are encouraged to determine the antecedents of other people’s emotions in novel situations or conflicts (“what happened to this person to make them feel this way?”).	0.333	3	0.634	0.353	0.956
ESS_76	Students are coached on how to express their feelings to others.	0.333	2.97	0.689	0.407	0.956

ESS_77	Students are shown how to appropriately express their emotions through modeling (e.g., teacher states, “I get angry when others laugh at me.”).	0.333	3.04	0.657	0.451	0.956
ESS_78	Students are encouraged to explore appropriate and inappropriate ways of how characters in a novel or story might express their emotions.	1	3.05	0.68	0.322	0.956
ESS_79	Students are encouraged to explore the most effective way to	1	3.17	0.6	0.509	0.956

	cope with their strong emotions.					
ESS_80	Students are treated with respect at all times.	1	3.55	0.534	0.524	0.956
ESS_81	Some students feel as if other students receive preferential treatment.	-0.333	2.6	0.671	-0.06	0.957
ESS_82	Students are comfortable telling the teacher when other students are disrespectful to them or interfere with their learning.	1	3.19	0.624	0.459	0.956
ESS_83	Interactions with students are always professional (e.g., appropriate boundaries, respectful language).	1	3.52	0.577	0.501	0.956

ESS_84	Students are encouraged to do their best each day.	-0.333	3.7	0.535	0.412	0.956
ESS_85	Students feel they are treated in the same way as their classmates (e.g., equal amount of attention; same consequences).	0.333	3.17	0.678	0.421	0.956
ESS_86	A warm and supportive environment is maintained for students in this class.	1	3.55	0.526	0.62	0.956
ESS_87	When students demonstrate prosocial behaviors (e.g., explaining an assignment to a peer), they are	0.333	3.41	0.57	0.507	0.956

	acknowledged or praised.					
ESS_88	If a student is struggling to make friends or cannot get along with other students, help or resources are provided (e.g., social skills training).	1	3.13	0.67	0.388	0.956
ESS_89	Students are encouraged to use appropriate problem-solving skills during conflicts with classmates.	1	3.35	0.589	0.573	0.956
ESS_90	Students receive positive feedback when they are appropriately assertive with their	0.333	3.2	0.591	0.415	0.956

	peers or stand up for themselves.					
ESS_91	Students are given examples of how to appropriately and inappropriately solve conflicts.	1	3.21	0.623	0.506	0.956
ESS_92	Teacher uses student conflicts as opportunities to model effective problem solving and communication strategies.	0.333	3.14	0.622	0.43	0.956
ESS_93	Students are provided with corrective feedback when they communicate in a disrespectful or aggressive way (e.g., using a rude tone of	0.333	3.29	0.553	0.57	0.956

	voice, using unkind words).					
AS_94	Students are given all the information they need to know in order to do well in this class (e.g. know when tests are scheduled, what material will be on tests, how to prepare for exams).	0.333	3.51	0.545	0.613	0.956
AS _95	Students are guided in creating study guides to help them prepare for exams.	0.333	3.11	0.763	0.377	0.956
AS _96	Students are shown or taught how to study for tests and quizzes.	0.333	3.21	0.684	0.482	0.956
AS _97	Students are aware of what material	1	3.52	0.593	0.598	0.956

	they will be tested on.					
AS _98	Students consolidate their notes from class and assigned readings to review the material for tests.	-0.333	2.97	0.779	0.355	0.956
AS _99	Students develop study aids (flashcards, practice exam questions) to prepare for upcoming tests and quizzes.	0.333	2.96	0.752	0.357	0.956
AS _100	Students are provided with learning objectives to ensure that they know what material is important.	1	3.43	0.54	0.568	0.956
AS _101	Students regularly receive homework	-0.333	3.04	0.849	0.294	0.956

	assignments to practice what has been learned in class.					
AS _102	Students complete a summative assessment (e.g., paper, project, test, quiz) after completing each academic unit or module.	0.333	3.32	0.612	0.47	0.956
AS _103	Students' knowledge of class content is evaluated on a consistent and timely basis.	1	3.39	0.567	0.586	0.956
AS _104	Not all students complete the same assignments. ®	0.333	2.37	0.713	-0.005	0.957
AS _105	Not all students take the same tests. ®	0.333	2.46	0.755	-0.034	0.957

AS _106	Test content directly reflects learning objectives.	1	3.47	0.527	0.632	0.956
AS _107	Students are assessed on the majority of the material that is covered in class.	0.333	3.39	0.542	0.604	0.956
AS _108	Test formats (e.g., short/long answer; multiple-choice) remain consistent across exams.	-0.333	3.15	0.734	0.3	0.956
AS _109	Unexpected assessments (e.g., “pop quizzes”) are administered in this class.®	0.333	2.65	0.931	-0.038	0.957
AS _110	Class activities or projects are used to gauge students’	1	3.3	0.59	0.499	0.956

	understanding of the material.					
AS _111	Students know how to improve their grades from the written feedback they receive.	1	3.3	0.614	0.467	0.956
AS _112	It is not possible to provide suggestions for improvement on all assignments that students turn in. ®	0.333	2.7	0.906	0.145	0.957
AS _113	It is not possible to provide students with timely feedback on assignments (e.g., homework, tests). ®	1	2.98	0.844	0.242	0.957
AS _114	Grading criteria and objectives for all assignments are clearly	1	3.4	0.593	0.645	0.956

	communicated to students					
AS _115	Students understand how assignments will be graded (e.g., scoring rubrics) prior to turning them in.	1	3.44	0.582	0.546	0.956
AS _116	Directions of how students can improve their grades are regularly communicated to caregivers.	0.333	3.2	0.58	0.52	0.956
AS _117	Students are provided with extra support when it is needed.	1	3.48	0.553	0.574	0.956
IT_118	Projects that require students to present course content in	0.333	3	0.645	0.306	0.956

	novel ways are regularly assigned.					
IT_119	Students are required to present material to their peers in a way that encourages active problem solving or critical thinking.	1	3.04	0.607	0.52	0.956
IT_120	Students are encouraged to question why they are learning certain material.	0.333	3.01	0.696	0.435	0.956
IT_121	Students are taught how to summarize the key concepts of book chapters or lessons.	1	3.28	0.574	0.438	0.956
IT_122	The majority of questions asked in	0.333	2.98	0.679	0.236	0.956

	this class are open-ended.					
IT_123	The majority of questions posed in this class are close-ended (e.g., require one-word answers). ®	-0.333	2.69	0.792	0.075	0.957
IT_124	Class discussions or debates limit the amount of material that can be covered in class	0.333	2.56	0.765	-0.043	0.957
IT_125	Questions are used as a way to stimulate class discussions.	1	3.38	0.54	0.524	0.956
IT_126	Students are given the opportunity to teach their peers or lead class lessons.	1	3.05	0.595	0.379	0.956
IT_127	A variety of methods (e.g., diagrams,	1	3.4	0.6	0.553	0.956

	videos, discussions, debates) are used to maintain student interest and to encourage participation.					
IT_128	The use of real-world examples is used to make the material more meaningful or relevant	1	3.46	0.576	0.587	0.956
IT_129	It is difficult to think of real-world examples that would resonate with students. ®	1	2.92	0.885	0.221	0.957
IT_130	It is difficult to find activities that keep students engaged. ®	1	2.87	0.863	0.264	0.957
IT_131	It is important to walk around the	-0.333	3.42	0.654	0.373	0.956

	classroom while lecturing to keep students attention. Students are expected to use outside resources (e.g.,					
IT_132	computer/internet, additional readings) to complete assignments or projects.	-0.333	2.93	0.746	0.299	0.956
IT_133	Both auditory and visual learning modalities (e.g., talking through a concept, drawing a flow chart) are used to help students understand new material.	1	3.39	0.541	0.565	0.956

IT_134	Students who need extra assistance are often paired with students who do well in the class when given group assignments.	0.333	3.14	0.622	0.319	0.956
IT_135	Group projects are assigned so students learn from one another.	1	3.12	0.686	0.392	0.956
IT_136	Students are taught memorization strategies (e.g., pneumonics) to aide in their learning of new material.	-1	3.05	0.661	0.319	0.956
IT_137	Previously learned content is worked into new lessons to provide repetition or	1	3.35	0.541	0.503	0.956

	to make connections across lessons. Experiential learning (e.g., games, field trips, experiments) is					
IT_138	used to present the material that is learned in class in a different way. Students' responses are repeated or expanded upon, so	1	3.15	0.702	0.448	0.956
IT_139	they feel heard and participation is encouraged.	0.333	3.35	0.54	0.553	0.956

APPENDIX J - ATEP Final Domain Names and Definitions

Emotional and Social Support: Teachers support and foster prosocial behaviors, such as appropriately processing and expressing emotions and using effective interpersonal skills.

Classroom Expectations and Routines: Teachers communicate behavioral and academic (e.g., how they will be assessed) expectations to students by making efforts to provide a stable, predictable classroom environment.

Assessment and Instructional Learning: teachers regularly assess students' mastery of topics covered in class while providing guidance to students on the use of effective study and critical thinking skills

Classroom Chaos: there is difficulty in consistency engaging in effective instructional techniques and stating and enforcing behavioral, academic, and classroom procedures.

Classroom Climate and Instructional Support: teachers create and maintain a classroom environment that aims to increase student engagement a supportive student-teacher relationship and classroom environment.

APPENDIX K - Descriptive Statistics of ATEP Final Item Pool – 58 Items

Five-factor structure of ATEP.

Item	CVR	Mean	Standard Deviation	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Factor 1 Emotional and Social Support					
73. Students are encouraged to take the emotional perspective of others (e.g., “what do you think this person is feeling?”).	.333	3.07	.654	.709	.892
76. Students are coached on how to express their feelings to others.	.333	3.03	.693	.728	.891
75. Students are encouraged to determine the antecedents of other	.333	3.04	.656	.690.	893

people's emotions in novel situations or conflicts ("what happened to this person to make them feel this way?").					
72. Students are taught how to label their emotions through reflection (e.g., "I can tell you are frustrated.").	1	2.96	.657	.652	.895
77. Students are shown how to appropriately express their emotions through modeling (e.g., teacher states, "I get angry when others laugh at me.").	.333	3.10	.670	.722	.891
79. Students are encouraged to	1	3.22	.603	.674	.894

explore the most effective way to cope with their strong emotions.					
91. Students are given examples of how to appropriately and inappropriately solve conflicts.	1	3.28	.625	.621	.897
88. If a student is struggling to make friends or cannot get along with other students, help or resources are provided (e.g., social skills training).	1	3.19	.680	.584	.898
74. Students feelings are normalized and validated (e.g., “Everyone gets angry.”).	.333	3.22	.611	.574	.899

78. Students are encouraged to explore appropriate and inappropriate ways of how characters in a novel or story might express their emotions.	1	3.09	.696	.350	.900
92. Teacher uses student conflicts as opportunities to model effective problem solving and communication strategies.	.333	3.19	.648	.547	.900
46. Students feel motivated to make good classroom decisions.	1	3.30	.570	.490	.902

Factor 2 - Classroom Expectations and Procedures

24. Behavioral expectations for this class are developmentally appropriate (e.g., raise hand to speak, follow instructions).	1	3.64	.546	.658	.887
10. Instructions for class and homework assignments are clearly outlined for students.	1	3.57	.540	.642	.887
20. The procedures for turning in class assignments or homework remain consistent.	.333	3.60	.555	.581	.890
21. Classroom rules are clearly stated to students.	1	3.72	.517	.629	.888
26. Students are provided with	1	3.49	.564	.665	.886

examples of appropriate behaviors (e.g., modeling desired behavior, class discussion).					
11. Academic expectations (e.g., meeting deadlines, completing assignments, studying for exams) set in this class are developmentally appropriate.	1	3.51	.551	.618	.889
27. Students in this class understand what behavior is expected of them.	.333	3.57	.560	.562	.891
50. Students have access to the materials they need.	1	3.52	.557	.634	.888

29. Students receive corrective feedback for inappropriate behaviors in a timely manner.	1	3.50	.571	.615	.889
54. Students are able to easily see the teacher and presented material during a class lesson.	1	3.61	.540	.562	.891
49. Students know where to find additional resources or supplies in the classroom (e.g., dictionaries, textbooks, pencils, paper).	.333	3.50	.571	.621	.888
59. Students are assigned a manageable amount of work each class.	-1	3.37	.549	.562	.892

Factor 3 - Learning and Instructional Techniques

95. Students are guided in creating study guides to help them prepare for exams.	.333	3.16	.776	.546	.897
111. Students know how to improve their grades from the written feedback they receive.	1	3.38	.620	.639	.891
96. Students are shown or taught how to study for tests and quizzes.	.333	3.27	.693	.634	.892
100. Students are provided with learning objectives to ensure that they know what material is important.	1	3.47	.543	.655	.891

116. Directions of how students can improve their grades are regularly communicated to caregivers.	.333	3.28	.604	.599	.893
115. Students understand how assignments will be graded (e.g., scoring rubrics) prior to turning them in.	1	3.48	.583	.657	.891
118. Projects that require students to present course content in novel ways are regularly assigned.	.333	3.08	.692	.441	.900
97. Students are aware of what material they will be tested on.	1	3.57	.586	.606	.893

114. Grading criteria and objectives for all assignments are clearly communicated to students.	1	3.47	.589	.665	.890
94. Students are given all the information they need to know in order to do well in this class (e.g. know when tests are scheduled, what material will be on tests, how to prepare for exams).	.333	3.55	.555	.603	.893
121. Students are taught how to summarize the key concepts of book chapters or lessons.	1	3.34	.580	.546	.895

103. Students' knowledge of class content is evaluated on a consistent and timely basis.	1	3.46	.563	.609	.893
119. Students are required to present material to their peers in a way that encourages active problem solving or critical thinking.	1	3.13	.642	.570	.894
106. Test content directly reflects learning objectives.	1	3.52	.530	.611	.893
Factor 4 - Classroom Chaos					
130. It is difficult to find activities that keep students engaged. ®	1	2.94	.912	.636	.852
129. It is difficult to think of real-world	1	3.00	.924	.668	.849

examples that would resonate with students. ®					
113. It is not possible to provide students with timely feedback on assignments (e.g., homework, tests). ®	1	3.05	.860	.586	.856
38. Students in this class receive more negative feedback than positive feedback about their behavior. ®	.333	3.18	.885	.632	.852
35. Students are unsure what will happen if they do not follow the rules. ®	1	3.30	.899	.569	.857
6. It is unclear what students are expected to do when they come to class. ®	.333	3.32	.886	.555	.859

66. Classroom routines and procedures are dependent on teacher-related factors (e.g., teacher energy, preparedness). ®	.333	2.66	.990	.479	.866
71. The classroom routine frequently requires adjustment. ®	1	2.57	.824	.472	.865
	Factor 5 - Classroom Climate				
139. Students' responses are repeated or expanded upon, so they feel heard and participation is encouraged.	.333	3.39	.546	.629	.842
125. Questions are used as a way to	1	3.44	.567	.612	.843

stimulate class discussions.					
137. Previously learned content is worked into new lessons to provide repetition or to make connections across lessons.	1	3.40	.555	.604	.844
128. The use of real-world examples is used to make the material more meaningful or relevant.	1	3.52	.570	.649	.840
127. A variety of methods (e.g., diagrams, videos, discussions, debates) are used to maintain student interest and to	1	3.47	.589	.559	.847

encourage participation.					
86. A warm and supportive environment is maintained for students in this class.	1	3.60	.513	.613	.843
87. When students demonstrate prosocial behaviors (e.g., explaining an assignment to a peer), they are acknowledged or praised.	.333	3.44	.567	.543	.849
134. Students who need extra assistance are often paired with students who do well in the class when given group assignments.	.333	3.19	.645	.412	.862

89. Students are encouraged to use appropriate problem-solving skills during conflicts with classmates.	1	3.41	.589	.571	.846
80. Students are treated with respect at all times.	1	3.61	.533	.522	.850

REFERENCES

Allen, K. P. (2010). Classroom management, bullying, and teacher practices. *Professional Educator*, *34*(1), n1.

Anderson, G. J., & Walberg, H. J. (1974). Learning environments. *Evaluating educational performance: A sourcebook of methods, instruments, and examples*, 81-98.

Arrindell, W. A., & Van der Ende, J. (1985). An empirical test of the utility of the observations-to-variables ratio in factor and components analysis. *Applied Psychological Measurement*, *9*(2), 165-178.

Baule, S. M., & Superior, W. I. (2020). The impact of positive behavior intervention support (PBIS) on suspensions by race and ethnicity in an urban school district. *AASA Journal of Scholarship & Practice*, *16*(4), 45-56.

Brandt, C., Mathers, C., Oliva, M., Brown-Sims, M., & Hess, J. (2007). Examining District Guidance to Schools on Teacher Evaluation Policies in the Midwest Region. Issues & Answers. REL 2007-No. 030. *Regional Educational Laboratory Midwest*.

Brophy, J. (1982, March). Research on the self-fulfilling prophecy and teacher expectations. In L. Shulman (Chair), *The self-fulfilling prophecy: Its origins and consequences in research and practice.* Symposium presented at the meeting of the American Educational Research Association, New York.

Brophy, J. E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of educational psychology*, *75*(5), 631.

Brophy, J. (1988). Educating teachers about managing classrooms and students. *Teaching and teacher Education*, *4*(1), 1-18.

Brophy, J. (2006). History of research on classroom management. *Handbook of classroom management: Research, practice, and contemporary issues*, 17-43.

Brophy, L. M., Jackson, M., & Crowe, S. F. (2009). Interference effects on commonly used memory tasks. *Archives of clinical neuropsychology*, *24*(1), 105-112.

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher education*, *16*(2), 239-253.

Cadavid, V., & Lunenburg, F. C. (1991). Locus of Control, Pupil Control Ideology, and Dimensions of Teacher Burnout.

Carretero-Dios, H., & Pérez, C. (2007). Standards for the development and review of instrumental studies: Considerations about test selection in psychological research. *International journal of clinical and health psychology*, *7*(3), 863-882.

Cattell, R. B. (1978). Fixing the number of factors: The most practicable psychometric procedures. In *The Scientific Use of Factor Analysis in Behavioral and Life Sciences* (pp. 72-91). Springer, Boston, MA.

Comrey, A. L., & Lee, H. B. (1992). A First Course in Factor Analysis, 2nd Edn. Hillsdale, NJ: L.

Conroy, M. A., Sutherland, K. S., Snyder, A., Al-Hendawi, M., & Vo, A. (2009). Creating a Positive Classroom Atmosphere: Teachers' Use of Effective Praise and Feedback. *Beyond Behavior*, *18*(2), 18-26.

Conroy, M. A., Sutherland, K. S., Snyder, A. L., & Marsh, S. (2008). Classwide interventions: Effective instruction makes a difference. *Teaching Exceptional Children*, *40*(6), 24-30.

Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational research methods*, *6*(2), 147-168.

Costello, A. B., & Osborne, J. (2005). Practical assessment, research & evaluation. *The Journal of Consumer Marketing*, *10*(7), 1-9.

Coughlin SS. Recall bias in epidemiologic studies. *J Clin Epidemiol.* 1990; 43: 87–91.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, *16*(3), 297-334.

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Dellinger, A.B., Bobbett, J.J., Olivier, D.F., & Ellett, C.D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. Teaching and Teacher Education, 24, 751–766. doi:10.1016/j.tate.2007.02.010

Downer, J. T., Booren, L. M., Lima, O. K., Luckner, A. E., & Pianta, R. C. (2010). The Individualized Classroom Assessment Scoring System (inCLASS): Preliminary reliability and validity of a system for observing preschoolers' competence in classroom interactions. *Early childhood research quarterly*, *25*(1), 1-16.

Doyle, W. (2009). Situated practice: A reflection on person-centered classroom management. *Theory into Practice*, *48*(2), 156-159.

Duke, D. L. (1979). Classroom Management. The Seventy-eighth Yearbook of the National Society for the Study of Education. Part II.

Emmer, E. T., & Evertson, C. M. (2016). *Classroom management for middle and high school teachers*. Pearson.

Emmer, E. T., & Evertson, C. M. (2017). *Classroom management for elementary school teachers*. Pearson.

Evertson, C. M., Emmer, E.T., Clements, B.S., Worsham, M.E., (1994). *Classroom management for elementary teachers*. Allyn & Bacon, A Division of Simon & Schuster, Inc., 160 Gould Street, Needham Heights, MA 02194.

Evertson, C. M., & Weinstein, C. S. (2006). Handbook of Classroom Management: Research. *Practice and Contemporary Issues. Mahwah NJ*.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, *4*(3), 272.

Fisher, D. & Fraser, B. (1983). Validity and use of the Classroom Environment Scale. *Educational Evaluation and Policy Analysis*, *5*(3), 261–271.

Fraser, C. O. (1980). Measurement in psychology. *British Journal of Psychology*, *71*(1), 23-34.

Fraser, B. J. (1982). Assessment of Learning Environments: Manual for Learning Environment Inventory (LEI) and My Class Inventory (MCI). Third Version.

Fraser, B. J. (1998). Classroom environment instruments: Development, validity and applications. *Learning environments research*, *1*(1), 7-34.

Fraser, B. J. (1990). *Individualized Classroom Environment Questionnaire*. Melbourne: Australian Council for Educational Research.

Fraser, B. J., McRobbie, C. J., & Giddings, G. J. (1993). Development and cross-national validation of a laboratory classroom environment instrument for senior high school science. Science Education, 77, 1-24.

Fraser, B. J., & O'Brien, P. (1985). Student and teacher perceptions of the environment of elementary school classrooms. *The Elementary School Journal*, *85*(5), 567-580.

Goe, L. (2007). The Link between Teacher Quality and Student Outcomes: A Research Synthesis. *National comprehensive center for teacher quality*.

Good, T. L. (1979). Teacher effectiveness in the elementary school. *Journal of teacher education*, *30*(2), 52-64.

Good, T. (1981). Teacher expectations and student perceptions: A decade of research. Educational Leadership, 38(5), 415-422.

Greenberg, J., Putman, H., & Walsh, K. (2014). Training Our Future Teachers: Classroom Management. Revised. *National Council on Teacher Quality*.

Gorsuch, R. L. (1983). Factor analysis. Lawrence Erlbaum. *Hillsdale, NJ*.

Gorsuch, R. L. (1990). Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research*, *25*(1), 33-39.

Guilford, J. P. (1954). Psychometric methods.

Hattie, J. (2003). Teachers Make a Difference, What is the research evidence?.

Hattie, J. (2009). The black box of tertiary assessment: An impending revolution. *Tertiary assessment & higher education student outcomes: Policy, practice & research*, 259-275.

Jackson, D. L. (2001). Sample size and number of parameter estimates in maximum likelihood confirmatory factor analysis: A Monte Carlo investigation. *Structural Equation Modeling*, *8*(2), 205-223.

Johnson, H. A., & Barrett, L. (2017). Your teaching strategy matters: how engagement impacts application in health information literacy instruction. *Journal of the*

Medical Library Association : JMLA, *105*(1), 44–48. https://doi.org/10.5195/jmla.2017.8

Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and social psychology review*, *9*(2), 131-155.

Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2011). Identifying effective classroom practices using student achievement data. *Journal of human Resources*, *46*(3), 587-613.

Kraft, M. A., & Blazar, D. L. (2013). Improving Teacher Practice: Experimental Evidence on Individualized Teacher Coaching. *Society for Research on Educational Effectiveness.*

Mantzicopoulos, P., French, B. F., Patrick, H., Watson, J. S., & Ahn, I. (2018). The stability of kindergarten teachers' effectiveness: a generalizability study comparing the framework for teaching and the classroom assessment scoring system. *Educational Assessment*, *23*(1), 24-46.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2016). *Applied multivariate research: Design and interpretation*. Sage publications.

Moos, R.H., & Moos, B.S. (1978). Classroom social climate and student absences and grades. *Journal of Educational Psychology, 70*(2), 263 – 269.

Oliver, R. M., Wehby, J. H., & Reschly, D. J. (2011). Teacher classroom management practices: Effects on disruptive or aggressive student behavior. *Campbell Systematic Reviews*, *7*(1), 1-55.

Petrasek, M., James, A., Noltemeyer, A., Green, J., & Palmer, K. (2021). Enhancing motivation and engagement within a PBIS framework. *Improving Schools*, 13654802211002299.

Pianta, R. C., K. M. La Paro, and B. K. Hamre. 2008. *Classroom Assessment Scoring System.*

Manual Pre-K. Baltimore, MD: Brookes

Rimm-Kaufman, S. E., Storm, M. D., Sawyer, B. E., Pianta, R. C., & LaParo, K. M. (2006). The Teacher Belief Q-Sort: A measure of teachers' priorities in relation to disciplinary practices, teaching practices, and beliefs about children. *Journal of School Psychology*, *44*(2), 141-165.

Rosenberg, M. S., & Jackman, L. A. (2003). Development, implementation, and sustainability of comprehensive school-wide behavior management systems. *Intervention in School and Clinic*, *39*(1), 10-21.

Rubie-Davies, C. M. (2007). Classroom interactions: Exploring the practices of high-and low-expectation teachers. *British Journal of Educational Psychology*, *77*(2), 289-306.

Rubie-Davies, C., Hattie, J., & Hamilton, R. (2006). Expecting the best for students: Teacher expectations and academic outcomes. British Journal of Educational Psychology, 76, 429–444. doi:10.1348/000709905X53589

Sandholtz, J. H. (2011). Preservice Teachers' Conceptions of Effective and Ineffective Teaching Practices. *Teacher Education Quarterly*, *38*(3), 27-47.

Sink, C. A., & Spencer, L. R. (2007). Teacher version of the My Class Inventory-Short Form: An accountability tool for elementary school counselors. *Professional School Counseling*, *11*(2), 2156759X0701100208.

Talari, K., & Goyal, M. (2020). Retrospective studies–utility and caveats. *JR Coll Physicians Edinb*, *50*, 398-402.

Velicer, W. F., & Fava, J. L. (1998). Affects of variable and subject sampling on factor pattern recovery. *Psychological methods*, *3*(2), 231.

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate behavioral research*, *25*(1), 1-28.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *The Journal of Educational Research*, *84*(1), 30-43.

Wang, J., Spalding, E., Odell, S. J., Klecka, C. L., & Lin, E. (2010). Bold ideas for improving teacher education and teaching: What we see, hear, and think.

Weinstein, C. S., Romano, M. E., & Mignano, Jr., A. J. (2011). Elementary classroom management: Lessons from research and practice (5th ed.). New York, NY, USA: McGraw Hill.

Wenglinsky, H. (2002). The link between teacher classroom practices and student academic performance. *Education policy analysis archives*, *10*, 12.

Whitehurst, G., Chingos, M. M., & Lindquist, K. M. (2014). Evaluating teachers with classroom observations. *Brown Center on Education Policy: Brookings Institute*.

Widaman, K. F. (2006). III. Missing data: What to do with or without them. *Monographs of the Society for Research in Child Development*, *71*(3), 42-64

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, *8*(3).

Yu, T., & Richardson, J. C. (2015). Examining reliability and validity of a Korean version of the community of inquiry instrument using exploratory and confirmatory factor analysis. *The Internet and Higher Education*, *25*, 45-52.